

> 华为ICT认证系列丛书

华为交换机学习指南

王达 主编









华为交换机学习指南

王达 主编





华为ICT认证系列丛书

华为交换机学习指南

王达　主编

人民邮电出版社

北京

序

当今世界的变化速度是史无前例的，我们的生活和工作的每一个方面都需要网络连接和信息获取。展望未来，ICT 仍然处于快速创新的阶段，移动性、云计算、大数据、社区化等ICT领域的新趋势，正在引领ICT行业开创新的格局；与此同时，现实世界也正在发生深刻的数字化变革，物联网、电子商务、数字媒体等正在促进传统产业不断地升级和重构。因此，以ICT产业为代表的数据世界和以传统产业为代表的物理世界的深度融合，将在不断驱动全球经济发展的同时，也将深刻地改变了人们的工作和生活，数字公民、数字企业和数字社会正在形成，必将引领新的ICT变革，也必将引领新的社会变革。

行业的发展离不开人才的支撑，产业的变革也将对ICT行业人才的知识体系和综合技能提出更高的挑战，“知识融合、技能跨界”将成为合格ICT人才的新标准。基于对未来趋势的把握，华为致力于培养优秀的ICT人才，通过华为ICT培训与认证体系希望帮助全社会消除数字鸿沟，确保人人享有信息、通讯的基本权利，促进ICT知识传递和效益提升，支撑ICT产业的可持续发展。

作为全球领先的信息与通信解决方案供应商，华为的产品与解决方案已广泛应用于金融、电力、能源、交通、企业、运营商、政府等各个行业。所以，华为与行业专家、高校老师合作编写了 “华为 ICT认证系列丛书”，旨在为广大用户、ICT从业者，以及愿意投身到ICT行业中的人士提供学习帮助。《华为交换机学习指南》就是其中一本，也是“华为ICT认证系列丛书”的第一本书。

《华为交换机学习指南》是我们与国内资深网络技术专家、业界知名作者——王达老师合作并出版的。这本书是从学习和实用的角度，基于学习的逻辑对知识点进行了系统的组织编排，书籍由浅入深，让读者逐步构建起系统的网络知识体系；同时该书在内容上注重理论和实践相结合，既有原理讲解，又有配置应用，让读者能够学以致用。希望王达老师的《华为交换机学习指南》以及后续的《华为路由器学习指南》能够帮助读者快速地学习华为产品技术，系统地建立网络知识体系，使读者在浩瀚的知识海洋中找到方向，不断提升，在ICT行业大展身手！

华为企业业务集团CEO
2013年11月

自序

本书出版背景

自从笔者出版了一些国际知名网络厂商的设备配置与管理方面的图书以来，一直有读者在追问我什么时候写本华为公司网络设备配置的图书。因为华为公司是国内、乃至世界范围内的网络通信设备领军企业，有着广泛的用户市场。然而在公开发行的图书市场中专门针对华为公司网络设备的图书却难觅踪影。其实笔者也的确早有这方面的考虑，所以从几年前开始就一直在留意、学习华为网络设备和技术。而正好就在2012年，有幸获得华为公司的信任，让笔者主编“华为ICT认证系列丛书”中的两本图书（后面还有一本《华为路由器学习指南》），最终促成了本书的顺利出版上市。在此要特别感谢华为技术有限公司的刘洋老师和人民邮出版社的王建军老师的高度信任与大力支持！

除了以上所说的用户需求外，本套丛书的推出还有一个大环境的需求。随着全球信息化的发展和重要性的提升，信息化程度已经成为国家发展的重要指标，国家都开始把提高民族网络通信产品和技术提上重要议程。所以大家开始更多地把目光聚焦于以华为公司为代表的网络通信领域民族品牌上，许多企业越来越重视采用本土网络通信设备，开发自己的核心技术，许多网络工程人员也越来越需要更全面地学习、使用本土产品。这也是催生本套丛书快速上市的一大重要因素。

服务与支持

本书得到了华为技术有限公司的许多专家的大力帮助与指导，并由他们进行全书内容的正确性和权威性审核，同时也得到了人民邮电出版社各位编辑老师的支持，在此代表全体编委成员一并表示最由衷的感谢！

由于编者水平有限，编写时间比较紧，因此尽管我们全体编写人员和出版社编辑老师花了大量时间和精力校验，但书中仍可能存在一些错误和瑕疵，敬请各位批评指正，万分感谢！大家可以通过以下渠道向我们反馈，提出宝贵意见。同时我们也将通过以下渠道为大家提供专业的服务：

1. 5个分片的超级QQ读者群（仅允许对应加入一个群）

华北地区（包括黑龙江、吉林、辽宁、内蒙古、北京、天津、河北）：101580747

华中地区（包括河南、山西、湖北、湖南、江西、安徽、贵州）：17201450

华东地区（包括上海、浙江、江苏、福建、山东、台湾）：32354930

华西地区（包括陕西、四川、重庆、宁夏、甘肃、青海、新疆、西藏）：54435786

华南地区（广东、广西、海南、云南、香港、澳门）：21576699

2. 3个专家博客

51CTO博客：http://winda.blog.51cto.com

CSDN博客：http://blog.csdn.net/lycb_gz

ChinaUnix博客：http://blog.chinaunix.net/uid/10659021.html

3. 2个认证微博

新浪微博：weibo.com/winda

腾讯微博：t.qq.com/winda2010

鸣谢

本书由王达主编并统稿，经过几十位编委、技术专家整整一年时间夜以继日地工作，一次次地严格审校、修改和完善，这本近千页巨作终于完成，并顺利高质量地出版上市。在此感谢华为技术有限公司各位专家慎密地技术审校和大力支持，感谢人民邮电出版社各位编辑老师，以及各位编委的辛勤工作！以下是

参与本书编写和技术审校人员名单。（排名不分先后）

编委人员：何艳辉、周健辉、何江林、李 想、郑达明、蔡学军、王 爽、黄丽君、余志坚、曾育文、罗广平、廖小妹、李 峰、胡海侨、罗巧芬、杨丽珍、陈玉生、刘胜华、朱碧霞、刘云根、卢翠环、陈妙娟、郑小建、罗裕玲、夏 强、谢桂安、黄高福

技术审校：陈 昊、刘立灿、莫 雯、周常青

前言

本书具有许多非常鲜明的特色：

● 本书是华为网络设备技能认证、培训的指定教材。

● 全国第一本，也是目前唯一的大型华为交换机配置与管理工具图书

本书所包括的内容非常全面、系统，从最基础的华为交换机设备选型、VRP系统访问和使用，配置文件的上传和下载，到主流应用的以太网接口、以太网链路聚合、iStack堆叠、CSS集群、各种VLAN划分、GVRP VLAN注册、VLAN聚合、VLAN映射、MUX VLAN、QinQ、QinQ 映射、VLAN 间路由、STP/RSTP/MSTP、ACL 和端口/VLAN/MAC地址镜像配置与管理，最后到高端应用的QoS、IGMP/IGMP Snooping/PIM/组播VLAN、基于MAC地址的安全管理、基于ARP的安全管理、AAA访问控制策略和802.1x认证、MAC地址认证和Portal认证配置与管理等无一不囊括其中。真正的“一册在手，别无所求”。

● 深入浅出的技术原理剖析与分层次配置示例完美结合

本书不仅有比较深入的各种华为交换机技术原理的剖析，而且列举了大量各种不同级别的应用配置示例。这种有机结合就可以使广大读者朋友，特别是初级读者朋友不再是孤立地学习这些枯燥的技术原理，而是能体验到这些技术原理在实际工作中的具体应用，反过来又加深了对这些技术原理的理解。另外，书中大量的配置示例也是分层次的，这样就使读者朋友不仅可以全面了解各具体配置命令的使用方法，更能深入地理解不同配置命令之间的相互关联及应用方法。

● 综合配置思路分析和详尽配置步骤介绍完美结合

本书在介绍华为交换机功能配置与管理时，注意配置思路分析与配置步骤介绍的完美结合，而不是机械地罗列出各种功能配置步骤。这样可使读者朋友在“知其然”的同时“知其所以然”，可以充分理解各种具体功能的基本配置和实现原理，可以在实际的网络设备配置工作中做到举一反三，灵活应用。

适用读者对象

本书的内容非常全面、系统，适合于各层次的读者，具体如下：

● 使用华为交换机产品的用户；

● 华为培训合作伙伴以及华为信息与网络技术学院的学员；

● 高等院校的计算机网络专业学生；

● 希望从零学习华为交换机配置与管理的读者；

● 以前没有系统学习过华为交换机配置与管理的读者；

● 看不懂华为交换机配置方案，没有掌握通用配置方法的读者；

● 希望有一本可在平时工作中查阅的大型华为交换机配置手册的读者。

本书介绍的交换机是已广泛应用于政府、金融、能源、交通、电力、教育、电信运营商等行业和企业市场。S9700 T比特核心路由交换机是面向下一代园区网核心和数据中心业务汇聚而专门设计开发的高端智能交换机, 在提供高性能的 L2/L3 层交换服务基础上，进一步融合了MPLS VPN、硬件 IPv6、桌面云、视频会议、无线等多种网络业务。S6700和S7700 万兆汇聚交换机广泛适用于园区网络、数据中心核心/汇聚节点，可对无线、话音、视频和数据融合网络进行控制。S5700 千兆接入交换机提供灵活的全千兆以太网接入。S2700和S3700百兆接入交换机为企业用户提供二层和三层的百兆接入能力。S1700 SMB交换机是为中小企业、网吧、酒店、学校等市场开发的新一代绿色节能以太接入交换机。

本书主要内容

本书是国内图书市场中第一本专门介绍华为交换机配置与管理的工具图书，也是华为ICT认证系列培训

教材。全书共分18章，近千页，各章基本内容如下。

第1章：全面介绍最新一代Sx700大系中各个系列（包括S1700、S2700、S3700、S5700、S6700、S7700和S9700系列）和S9300系列交换机产品的主要特点、各机型硬件配置和软件功能特性，以及主要应用。

第2章：全面介绍新一代Sx700大系中所使用的VRP5.x系统的基础知识和基本使用方法，包括VRP CLI的使用和视图，以及VRP系统软件、VRP系统配置文件和VRP文件系统管理等。

第3章：全面介绍用户界面（包括Console用户界面和VTY用户界面），用户/命令级别、Console本地登录，Telnet、STelnet、HTTP和HTTPS远程登录，以及FTP、SFTP、SCP和FTPS远程文件管理配置方法。

第4章：全面介绍以太网接口编号规则，Eth-Trunk和E-Trunk链路聚合原理，基本参数、接口属性、端口组、端口隔离、手工配置Eth-Trunk、LACP Eth-Trunk、Eth-Trunk子接口、Loopback接口、Null接口和E-Trunk配置与管理。

第5章：全面介绍iStack堆叠和CSS集群工作原理，以及堆叠卡/业务口iSatck堆叠，集群卡/业务口CSS集群配置与管理。

第6章：全面介绍VLAN基础知识、二层以太网接口类型及各自的数据收发规则和GVRP VLAN注册原理，以及基于端口/基于MAC地址/基于 IP子网/基于 IP协议/基于策略的VLAN划分、GVRP VLAN注册、各种VLAN间路由方案（包括VLANIF接口方案、以太网子接口方案和VLAN交换方案）和管理VLAN的配置与管理。

第 7章：全面介绍VLAN聚合、MUX VLAN、基本QinQ、灵活QinQ、QinQ映射和VLAN映射工作原理及配置与管理。

第8章：全面介绍STP、RSTP和MSTP基础知识和技术原理及配置与管理。

第9章：全面介绍各种ACL类型（基本ACL、高级ACL、二层ACL、用户自定义ACL、自反ACL）的基础知识和各自的配置与管理方法，以及ACL在简化QoS流策略中的应用配置。

第10章：全面介绍与QoS有关的基础知识和技术原理，包括各种QoS优先级及优先级映射和QoS流策略基础知识，以及各种流量监管、流量整形、拥塞避免和拥塞管理 技术原理。

第11章：全面介绍各种QoS优先级映射、流量监管、流量整形、拥塞避免、拥塞管理、复杂QoS流策略的配置与管理。

第12章：全面介绍在IPv4网络中应用的IGMP、PIM（包括PIM-DM和PIM-SM两种模式）、MSDP三层IP组播协议，IGMP snooping、组播VLAN二层 IP组播协议的基础知识和各自的工作原理，以及组播路由管理原理。

第 13章：全面介绍在 IPv4组播网络中主要应用的 IGMP、PIM、IGMP snooping、组播VLAN协议功能的配置与管理。

第14章：全面介绍端口镜像、VLAN镜像和MAC地址镜像原理和各自的配置与管理。

第15章：全面介绍静态MAC地址表项、端口安全（包括动态安全MAC地址和Sticky MAC地址）、MAC地址防漂移、MAC地址漂移检测、MAC-spoofing-defend、端口桥等功能的配置与管理。

第16章：全面介绍防御ARP泛洪攻击（包括ARP报文限速、ARP表项严格学习、ARP表项限制、ARP Miss消息抑制和ARP表项老化等功能）和防ARP欺骗攻击（包括ARP表项固化、动态ARP检测、ARP防网关冲突、免费ARP报文发送、MAC地址一致性检查和ARP报文合法性检查等）的配置与管理方法。

第17章：全面介绍本地方式、RADIUS服务器方式和HWTACACS服务器方式AAA访问控制方案（包括认证、授权和计费方案）基础知识及配置与管理方法。

第18章：全面介绍802.1x认证、MAC地址认证和Portal认证基础知识及配置与管理方法。

阅读注意地方

在阅读本书时，请注意以下几个地方：

● 书中所有讲到的“S系列”均仅指最新的Sx700大系和S9300/9300E系列产品。

● 在不同的VRP系统版本中存在Quidway和HUAWEI两种缺省主机名。

● 为了避免内容重复，与华为路由器相同的功能部分，如 DHCP、DNS、各种路由协议、VRRP、VPN，本书均没有介绍，请参见后面即将出版的《华为路由器学习指南》一书。

● 书中的配置代码中，粗体字部分是命令本身或关键字选项部分，是不可变的；斜体字部分是命令或者关键字参数部分，是可变的。

● 在介绍各种交换机技术及功能配置说明过程中，对于一些需要特别注意的地方均以粗体字格式加以强调，以便读者在阅读学习时引起特别注意。

● 在介绍各种功能配置的过程中针对不同S系列交换机中相同功能的不同配置方法或参数取值范围做了特别说明，以便读者能全面了解不同系列交换机的不同配置方法和参数取值范围。

● 在介绍各种功能特性时明确列出各个S系列交换机对这些特性的支持情况，以便读者明确了解自己所使用的机型对相应特性的支持情况。

目　录

第1章　华为园区交换机的选型和应用

华为公司园区交换机就是其S系列交换机，目前最新一代是Sx700系列。它包括S1700、S2700、S3700、S5700、S6700、S7700和S9700七大系列，另外S9300系列交换机也是目前主流高端交换机系列，可全方位满足于个人/小型企业用户，中小型和大中型企业园区网，以及各种规模数据中心的不同网络环境、不同应用场景的客户需求。

Sx700系列新一代园区交换机有一个共同的特点，那就是绿色节能，采用了新型的低噪声、低辐射、空闲端口休眠等自动节能技术的高集成芯片电路设计。在工作层次方面，从普通的二层以太网交换机到三层交换机，再到多层路由交换机，支持二至四层各种功能特性；在端口带宽方面，从百兆到千兆、万兆，再到最新的10万兆全面覆盖；在交换容量方面从几十Gbit/s到几十Tbit/s也全面覆盖，可满足所有企业网、园区网和数据中心用户的各方面硬件配置、交换/路由功能和交换性能需求。

本章将重点介绍各大系列的主要特点、机型规格、主要特性及主要应用，同时将对一些功能和应用场景类似的交换机系列，如S7700系列、S9300系列和S9700系列进行横向综合比较，以便更好地帮助用户对华为园区交换机的选型。

1.1 华为园区交换机基础

华为公司的交换机产品线非常齐全，从大的分类来看主要分传统以太网交换机和当前热门的云计算交换机，本书仅介绍传统以太网交换机，也就是华为的园区S系列交换机。

1.1.1 华为园区交换机概述

华为园区交换机是针对各种企业园区网而开发的以太网交换机产品系列，又称S系列交换机。目前最新的 S 列是 Sx700 系列，其中的“7”代表产品大系号，“x”代表不同的产品系列，包括S1700、S2700、S3700、S5700、S6700、S7700和S9700共7个产品系列，是华为公司自主研发的新一代绿色、节能型交换机系列产品。目前市场上仍主流应用的还有S9300高端T比特路由交换机。

这些主流的华为S系列交换机的软硬件配置、功能，以及性能档次各不相同，定位于不同的用户，可全面满足于家庭网络、中小型企业园区网、大型企业园区网，以及各种规模的数据中心等各种网络环境下，对交换机功能、性能、业务处理、安全和管理等各方面的需求。它们的基本分类或各系列的主要功能版本如图1-1所示。

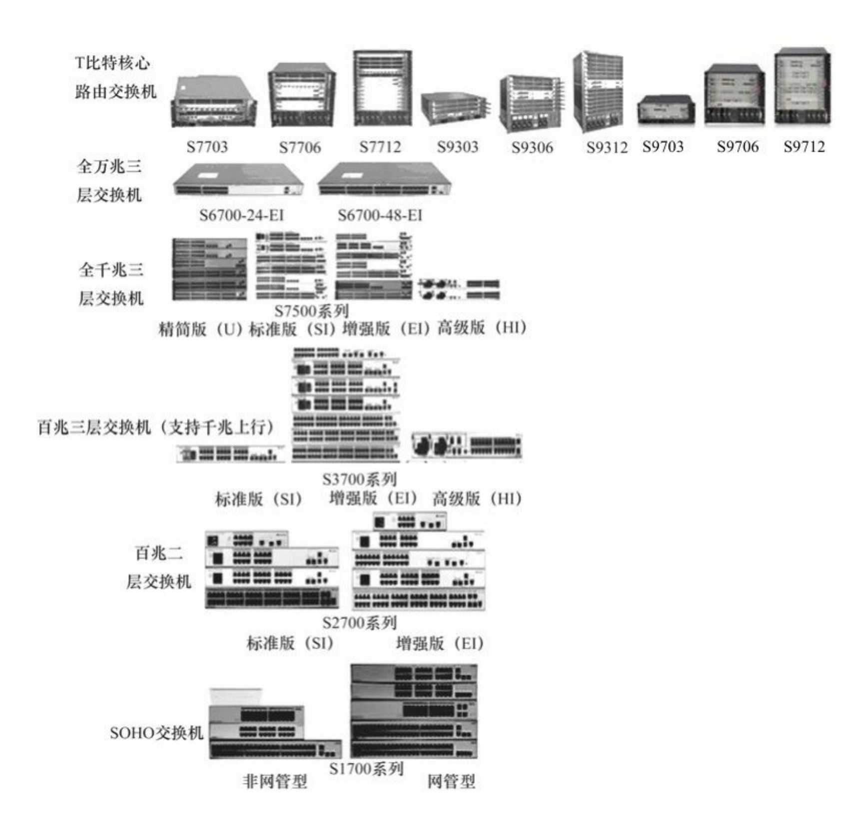


图1-1 华为S系列园区交换机的基本分类

从以上基本分类可以看出，S1700和S2700两大系列都属于百兆（部分机型提供了千兆端口）二层交换机，主要定位于中小型网络的接入层，实现百兆桌面接入；S3700及以上各大系列均属于三层交换机系列，其中S3700系列属于百兆三层交换机（支持千兆上行），主要定位于中型网络的接入层和小型网络的汇聚层；S5700系列属于全千兆三层交换机（支持万兆上行），主要定位于大中型网络的汇聚层，实现多业务的汇聚与转发，以及中小型数据中心的接入层，实现服务器的千兆接入；S6700 系列属于万兆三层交换机，主要定位于中型网络的核心层、大中型网络的汇聚层，实现多业务的汇聚与转发，以及大中型数据中心的接入层，实现高性能服务器的万兆接入；S7700、S9300和S9700系列都属于T比特路由交换机，提供极高性能的业务数据处理和路由性能，借助于CSS （Cluster Switch System，集群交换系统）实现业界最高性能的交换机集群，同时消除单点故障，主要定位于大型网络的核心层和大型数据中心的核心层。同一系列交换机也可能分为“精简版”（LI）、“标准版”（SI）、“增强版”（EI）和“高级版”（HI）等不同版本。需要注意的是，这里的版本不是指交换机的软件版本，而是指不同的硬件型号。

由于应用场景和产品定位不同，华为新一代Sx700系列交换机各子系列支持的功能有所差异。但各子系列交换机均采用统一的VRP平台，这使得相同的功能在不同的产品上配置基本相同，这样用户只需要关注产品的特性差异，而不需要针对不同的产品学习不同的配置方法。需要注意的是S1700与其他系列不同，它定位于SOHO级应用，功能比较简单，一般不需要配置或者通过简单的Web配置即可应用。

1.1.2 华为园区交换机的用户定位

图1-1是从交换机功能和性能方面对华为S系列园区交换机的基本分类，还可以从交换机的主要应用环境或用户定位来划分。目前一般的企业园区网接入层主要应用的是S2700和S3700两大系列，汇聚层主要应用的是S5700系列，核心层主要应用的是S7700、S9300和S9700系列，数据中心接入层主要应用的是 S5700

和 S6700 系列，数据中心核心层主要应用的是S7700、S9300和S9700系列，具体如下。大家可以根据这些交换机的基本用户定位进行华为园区交换机的选型。各系列交换机的主要性能和硬件配置将在本章后面具体介绍。

1．数据中心交换机

根据数据中心网络规模大小和性能要求的高低，数据中心的核心层可以采用S9700、S9300或S7700系列交换机，接入层可以采用S6700或S5700系列交换机。其中S9700、S9300和S7700三大系列均是T比特级的核心路由交换机，不仅单端口带宽非常高（如S9700系列最高可达40Gbit/s），而且交换能力强（如S9700系列最大可达18.56Tbit/s交换容量），可满足最苛刻的大型数据中心对端口带宽和交换性能的要求。接入层的S6700和S5700系列分别是全万兆和全千兆高性能三层交换机，最大交换容量分别可达960Gbit/s和256Gbit/s，可全面满足中、小型数据中心对端口带宽和交换性能的要求。随着技术的发展，华为公司还将持续推出更高性能的交换机，不断满足数据中心用户对交换机端口带宽和交换性能日益增长的需求。

2．核心交换机

根据园区网络规模大小和性能要求的高低，核心层可以分别采用 S9700、S9300、S7700、S6700、S5700和S3700系列交换机。其中S9700和S9300系列可作为大型园区网络的核心交换机；S7700、S6700 系列可作为中型园区网络的核心交换机；S5700 和 S3700系列可分别作为中、小型园区网络的核心交换机。

3．汇聚交换机

根据园区网络规模大小和性能要求的高低，汇聚层可以分别采用 S7700、S6700、S5700和S3700系列交换机。其中S7700和S6700系列可作为大型园区网络的汇聚交换机；S5700系列可作为中型园区网络的汇聚交换机，S3700系列可作为中、小型园区网络的汇聚交换机。

4．接入交换机

根据园区网络规模大小和性能要求的高低，接入层可以分别采用 S5700、S3700、S2700和S1700系列交换机。其中S5700系列可作为大型园区网络的全千兆接入交换机；S3700和S2700系列可作为中小型园区网络的百兆到桌面（支持千兆上行）的接入交换机，S1700系列可作为小型园区网络的百兆到桌面（部分支持千兆上行）接入交换机。

1.1.3 华为园区交换机的命名规则

了解设备的命名规则对于一个专业的网络工程人员来说既是非常重要的，也是非常必要的，因为这样就可以直接从设备名称中获知比较全面的硬件配置信息，以确定所需选择的机型，而不需要四处查找每一机型的具体配置资料。

在华为 S 系列园区交换机中，每个系列中都有许多种不同机型，特别是像 S1700、S2700、S3700、S5700这样应用范围比较广、机型比较多的中低端产品系列，每个系列中的每款机型的硬件配置或多或少有所不同，所以在正式介绍这些交换机系列前先介绍它们的命名规则，以便能快速地进行华为S系列交换机选型。但因为不同系列的主要应用环境和所包括的机型各不相同，所以它们在命名规则上也存在许多不同。下面分别予以介绍。

1．S1700系列机型的命名规则

S1700系列比较特殊，它是专门为个人和小型企业用户量身打造的SOHO级交换机。 由于应用、功能比较简单，一般不需要配置或者通过简单的Web配置即可使用。目前S1700系列中，网管型和非网管型交换机各有5款机型，具体将在本章后面介绍。下面以S1700-8-AC、S1700-28GFR-4P-AC和S1700-52FR-2T2P-AC 3款机型为例介绍S1700系列交换机的命名规则，如图 1-2 所示。各部分含义说明如表1-1所示。

S1700-8-AC
A B C I

S1700-28GFR-4P-AC
C DEFGH I

S1700-52FR-2T2P-AC
C EFGHGH I

图1-2 S1700系列交换机的命名规则

表1-1 S1700系列命名规则中各部分的含义

标号	含义
A	表示设备为交换机
B	表示产品系列，其中“17”表示 17 系列，“00”是子系列号
C	表示最大可用端口数。S1700 系列交换机支持的最大端口数不同，目前分别为 8、24、28、52 个
D	表示下行端口类型，G 为千兆端口。如果无此部分则表示该机型的下行端口为百兆端口
E	表示设备的网管类型，F 代表全网管类型，W 代表 Web 网管型。无此部分表示该机型为非网管机型
F	表示设备安装结构，R 表示为机架型结构。无此部分表示该机型为桌面型结构
G	表示上行端口数
H	表示对应的上行端口的类型，其中 T 表示双绞线以太网电口，P 表示 SFP 模块光口，TP 表示光口和电口的 Combo 端口，无此部分表示该机型无上行端口
I	表示设备的供电方式，目前 S1700 系列仅支持交流供电方式（AC）

2．S2700系列机型的命名规则

为了满足不同用户的需求，S2700 系列提供了多款机型。下面以 S2700-26TP-PWR-EI、S2710-52P-SI-AC、S2700-52P-EI-AC和S2700-9TP-SI为例介绍 S2700 系列交换机的命名规则，如图 1-3所示。各部分的具体含义如表1-2所示。

S2700-26TP-PWR-EI
AB C D E F G

S2710-52P-SI-AC
AB C D E G H

S2700-52P-EI-AC
D E G H

S2700-9TP-SI
D E G

图1-3 S2700系列交换机的命名规则

表1-2 S2700系列交换机命名规则中各部分的含义

标号	含义
A	表示设备为交换机
B	表示产品系列，其中“27”表示 27 系列
C	表示产品不同子系列
D	表示最大可用端口数。S2700 系列交换机支持的最大端口数不同，目前分别为 9、18、26、52 个
E	表示上行端口的类型，其中 TP 表示上行端口为支持光口和电口的 Combo 口，P 表示上行端口为光口
F	表示设备支持 PoE 供电，无此部分表示该机型不支持 PoE 供电
G	表示设备软件版本类型，其中 EI 表示设备为增强版本，包含某些高级特性；SI 表示设备为基本版本，包含基础特性
H	表示设备的供电方式，其中 AC 表示设备为交流供电，DC 表示设备为直流供电

3．S3700系列交换机的命名规则

同样，为了满足不同用户的需求，S3700系列提供了多款机型，用户可以根据不同的网络需求进行灵活

的选择。下面以S3700-28TP-PWR-EI、S3700-52P-EI-24S-DC、S3700-28TP-EI-MC-AC和S3700-28TP-SI-AC为例介绍S3700系列交换机的命名规则，如图1-4所示。各部分的具体含义如表1-3所示。

S3700-28TP-PWR-EI
A B C D E F

S3700-52P-EI-24S-DC
C D F G H

S3700-28TP-EI-MC-AC
C D F I H

S3700-28TP-SI-AC
C D F H

图1-4 S3700系列交换机的命名规则

表1-3 S3700系列交换机命名规则中各部分的含义

标号	含义
A	表示设备为交换机
B	表示产品系列，其中“37”表示37系列
C	表示最大可用端口数。S3700系列交换机支持的最大端口数不同，目前分别为28、52个
D	表示上行端口的类型，其中P表示上行端口为光口；TP表示上行端口为支持光口和电口的Combo口
E	表示设备支持PoE供电。无此部分表示该机型不支持PoE供电
F	表示设备软件版本类型，其中HI表示设备是高级版本，包含高性能OAM、内置RTC时钟等特性；EI表示设备为增强版本，包含某些高级特性；SI表示设备为基本版本，包含基础特性
G	表示下行端口的类型，24S表示S3752P-EI-24S的24个下行接口为光口。无此部分表示该机型中所有的下行接口均为电口
H	表示设备的供电方式，其中AC表示设备为交流供电，DC表示设备为直流供电
I	表示设备具有监控口，无此部分表示该机型无监控口

4．S5700系列交换机的命名规则

在所有Sx700系列中，S5700属于中间档次，应用范围最广，所以它的机型非常多，划分了从高到低的HI、EI、SI和 LI 4个功能版本系列，用户可以根据不同的网络需求进行灵活的选择。下面以S5710-28C-EI、S5700S-52P-LI-AC、S5700-48TP-PWR-SI、S5700-28C-EI-24S 和 S5700-28C-HI为例，介绍S5700系列交换机的命名规则，如图1-5所示。各部分的具体含义如表1-4所示。

S5710-28C-EI
A B C E F H

S5700S-52P-LI-AC
A B C D E F H J

S5700-48TP-PWR-SI
E G H

S5700-28C-EI-24S
I

S5700-28C-HI
H

图1-5 S5700系列交换机的命名规则

表1-4 S5700系列交换机命名规则中各部分的含义

标号	含义
A	表示设备为交换机
B	表示产品系列，其中“57”表示 57 系列
C	表示产品不同子系列
D	S 表示该机型为商业型号
E	表示最大可用端口数。S5700 系列交换机支持的最大端口数不同，目前分别为 24、28、48、52 个
F	表示上行端口的类型，其中 C 表示设备支持插卡，上行端口为 2，4 或 8；TP 表示上行端口为支持光口和电口的 Combo 口；P 表示上行端口为光口
G	表示设备支持 PoE 供电。无此部分表示该机型不支持 PoE 供电
H	表示设备软件版本类型，其中 EI 表示设备为增强版本，包含某些高级特性；SI 表示设备为基本版本，包含基础特性；HI 表示设备是高级版本，包含高性能 OAM、内置 RTC 时钟等特性；LI 表示设备是简化版本
I	表示下行接口的类型，24S 表示 S5700-28C-EI-24S 的 24 个下行接口为光口。无此部分表示该机型中所有的下行接口均为电口
J	表示设备的供电方式，其中 AC 表示设备为交流供电，DC 表示设备为直流供电

5．S6700系列交换机命名规则

S6700 系列属于中高端的万兆以太网交换机系列，由于应用环境比较单一，所以它的机型很少，目前仅有S6700-24-EI和S6700-48-EI两款机型。因为机型比较少，各机型的功能极其相似，所以它们的命名规则与前面介绍的各系列不一样。下面以S6700-48-EI为例介绍S6700设备的命名规则，如图1-6所示，各部分的具体含义如表1-5所示。

S6700-48-EI
A B C D

图1-6 S6700系列交换机的命名规则

表1-5 S6700系列交换机命名规则中各部分的含义

标号	含义
A	表示设备为交换机
B	表示产品系列，其中“67”表示 67 系列
C	表示最大可用端口数。S6700 系列设备支持的最大端口数不同，目前分别为 24、48 个
D	表示设备软件版本类型，其中 EI 表示设备为增强版本，包含某些高级特性

6．S7700/9300/9700系列交换机的命名规则

在华为公司S系列园区交换机中，S7700、S9300和S9700系列都是机箱式结构的交换机系列，属于高端交换机系列。它们主要应用于城域网中的业务接入、汇聚和传输层，作为城域网的接入和汇聚节点。机箱式结构可以十分方便地通过插入板卡进行端口或功能上的扩展，它们的命名规则也主要是根据所配置的插槽数来区分的，很简单。下面仅以S7706为例介绍S7700/9300/9700系列交换机的命名规则，如图1-7所示，各部分的具体含义如表1-6所示。

S7706
A B C

图1-7 S7700/9300/9700系列交换机的命名规则

表1-6 S7700/9300/9700系列交换机命名规则中各部分的含义

标号	含义
A	表示设备为交换机
B	表示产品系列，其中“77”表示 77 系列，“93”代表 93 系列，“97”代表 97 系列
C	表示提供的业务板槽位数，目前分别为 3、6、12 个

1.2 S1700系列交换机的选型与应用

S1700 系列企业交换机是华为公司专门针对个人用户和小型企业推出的新一代绿色节能二层以太网接入交换机，广泛应用于中小企业、网吧、酒店、学校等以太接入场景。S1700系列交换机分为网管型和非网管型两大类。由于功能简单，所以其网管型交换机可通过Web或者SNMP方式进行配置与管理，而不用通过CLI（命令行界面）以命令方式进行配置。

1.2.1 S1700系列机型及基本配置

S1700系列目前主要有10款机型，其中网管型和非网管型各5款，它们的基本配置如表1-7所示。

表1-7 S1700系列机型及基本配置

管理方式	型号	产品外观	端口数	基本配置
非网管型	S1700-8-AC		8	• 8 个 10/100Mbit/s 自适应以太网电口 • 交流供电 • 包转发率：1.2Mpps • 交换容量：1.6Gbit/s
	S1700-24-AC		24	• 24 个 10/100Mbit/s 自适应以太网电口 • 交流供电 • 包转发率：3.6Mpps • 交换容量：4.8Gbit/s

（续表）

管理方式	型号	产品外观	端口数	基本配置
非网管型	S1700-52R-2T2P-AC		52	• 48 个 10/100Mbit/s 自适应以太网电口，支持两个 GE 电口和两个 GE 光口 • 交流供电 • 包转发率：13.2Mpps • 交换容量：17.6Gbit/s
	S1700-8G-AC		8	• 8 个 10/100/1000Mbit/s 自适应以太网电口 • 交流供电 • 包转发率：12Mpps • 交换容量：16Gbit/s
	S1724G-AC		24	• 24 个 10/100/1000Mbit/s 自适应以太网电口 • 交流供电 • 包转发率：36Mpps • 交换容量：48Gbit/s
Web 网管型	S1728GWR-4P-AC		28	• 24 个 10/100/1000Mbit/s 自适应以太网电口，支持 4 个 GE SFP 独立光口 • 交流供电 • 包转发率：42Mpps • 交换容量：56Gbit/s
SNMP 网管型	S1700-28FR-2T2P-AC		28	• 24 个 10/100Mbit/s 自适应以太网电口，支持两个 GE 电口和两个 GE SFP 光口 • 交流供电 • 包转发率：9.6Mpps • 交换容量：12.8Gbit/s
	S1700-52FR-2T2P-AC		52	• 48 个 10/100Mbit/s 自适应以太网电口，支持两个 GE 电口和两个 GE SFP 光口 • 交流供电 • 包转发率：13.2Mpps • 交换容量：17.6Gbit/s
	S1700-28GFR-4P-AC		28	• 24 个 10/100/1000Mbit/s 自适应以太网电口，支持 4 个 GE SFP 光口 • 交流供电 • 包转发率：42Mpps • 交换容量：56Gbit/s
	S1700-52GFR-4P-AC		52	• 48 个 10/100/1000Mbit/s 自适应以太网电口，支持 4 个 GE SFP 光口 • 交流供电 • 包转发率：78Mpps • 交换容量：104Gbit/s

1.2.2 S1700系列交换机的规格

在 S1700 系列的 5 款非网管型台交换机中， S1700-8-AC、S1700-24-AC 提供10/100Base-T以太网电接口；S1700-8G-AC和S1724G-AC提供10/100/1000Base-T以太网电接口；S1700-52R-2T2P-AC 提供 48 个 10/100Base-T 以太网电接口、两个10/100/1000Base-T以太网电接口和两个1000Base-X以太网光接口；5款网管型交换机均提供10/100/1000Base-T以太网电口和1000Base-X以太网光口选择，同时支持Access、Trunk和Hybrid等多种接口类型，以及STP/RSTP和MSTP等生成树技术支持。

对于千兆光纤连接，非网管型的S1700-52R-2T2P-AC和5款网管型S1700系列交换机均提供可插拔的SFP（Small Form-Factor Pluggable，小型可插拔）类型光模块，光纤长度可以根据用户对传输距离的需求灵活选配。非网管型S1700系列交换机的规格如表1-8所示，网管型S1700系列交换机的规格如表1-9所示。

表1-8 非网管型S1700系列交换机的规格

型号	S1700-8-AC	S1700-24-AC	S1700-52R-2T2P-AC	S1700-8G-AC	S1724G-AC
固定端口	8 个 10/100Mbit/s 电口	24 个 10/100Mbit/s 电口	48 个 10/100Mbit/s 电口	8 个 10/100/1000Mbit/s 电口	24 个 10/100/1000Mbit/s 电口
可选端口	无	无	两个 GE 电口，两个 GE SFP 光口	无	无
MAC 地址表	8kB （1k=1024）	8k	8k	8k	8k
EEE	不支持	不支持	不支持	不支持	支持

表1-9 网管型S1700系列交换机的规格

型号	S1700-28FR-2T2P-AC	S1700-52FR-2T2P-AC	S1700-28GFR-4P-AC	S1700-52GFR-4P-AC	S1728GWR-4P-AC
固定端口	24 个 10/100Mbit/s 电口	48 个 10/100Mbit/s 电口	24 个 10/100/1000Mbit/s 电口	48 个 10/100/1000Mbit/s 电口	24 个 10/100/1000Mbit/s 电口
可选端口	两个 GE 电口，两个 GE SFP 光口	两个 GE 电口，两个 GE SFP 光口	支持 4 个 GE SFP 光口	支持 4 个 GE SFP 光口	支持 4 个 GE SFP 光口
MAC 地址表	8k	8k	8k	8k	8k
EEE	不支持	不支持	支持	支持	支持

1.2.3 S1700网管型交换机的主要特性

尽管S1700系列交换机不支持VRP系统，但其中的网管型交换机仍可通过Web方式提供可全面满足中小型企业接入层交换机所需求的功能特性，如支持基本的VLAN划分、STP/RSTP；支持Access、Trunk和Hybrid端口类型；支持端口汇聚、IGMP Snooping组播、QoS、802.1x认证和AAA访问控制策略等。S1700网管型交换机的主要特性如表1-10所示。

表1-10 S1700网管型交换机的主要特性

网管类型	Web 网管型交换机 （仅 S1728GWR-4P-AC 机型）特性	SNMP 网管型交换机特性
VLAN	• 支持 256 个 VLAN • 支持 Access、Trunk 和 Hybrid 端口类型 • 支持管理 VLAN 和 Voice VLAN	• 支持 4k（1k =1024，4k =4096）个 VLAN • 支持 Access、Trunk 和 Hybrid 端口类型 • 支持管理 VLAN 和 Voice VLAN
STP	支持 STP（IEEE 802.1d）和 RSTP（IEEE 802.1w）	支持 STP、RSTP 和 MSTP（IEEE 802.1s）
端口汇聚	• 支持 12 组汇聚组，每组最多 8 个端口 •支持静态 LACP	
端口镜像	支持基于端口的双向流量镜像	• 支持基于端口的双向流量镜像 • 镜像端口支持 Trunk 类型
端口带宽控制	支持对出入端口的报文流量进行限速，粒度最小为 64kbit/s	
广播风暴抑制	• 支持基于端口速率的风暴抑制 • 支持端口流量达到风暴抑制门限时发送告警	
组播	支持 IGMP Snooping，最多支持 256 个组播组	• 支持 IGMP Snooping，最多支持 256 个组播组 • 支持用户快速离开机制
QoS	• 支持绝对优先级、WRR 两种调度方式 • 支持每端口 4 个队列 • 支持根据 802.1p/DSCP 队列调度	• 支持绝对优先级、WRR 两种调度方式 • 支持每端口 8 个队列 • 支持根据 802.1p/DSCP 队列调度
安全特性	• 支持基于端口的 MAC 过滤，以及 802.1x 认证和 RADIUS 认证 • 支持端口隔离	• 支持硬件 ACL • 支持基于端口的 MAC 过滤，以及 MAC 认证、802.1x 认证和 RADIUS 认证 • 支持端口隔离和风暴抑制 • 支持系统自防御，防止广播流、ARP、ICMP、TCP、虫病毒、DOS 等攻击 CPU • 支持 DHCP Snooping
设备管理	• 支持 Web 网管 • 支持 DHCP-client • 支持一键还原	• 支持 SNMP、Web 网管（支持 HTTPS）方式 • 支持 DHCP-client • 支持用户口令保护 • 支持一键还原
设备维护	• 支持 Syslog（系统日志） • 支持 Ping 检测和 VCT（Virtual Cable Test，虚拟电缆检测） • 支持链路层发现协议 LLDP（Link Layer Discovery Protocol）	• 支持 RMON（Remote Network Monitoring，远程网络监控） • 支持 Syslog • 支持 Ping 检测/Traceroute 和 VCT • 支持链路层发现协议 LLDP

1.2.4 S1700系列交换机的主要应用

S1700 系列交换机是专门为中小型企业、酒店、学校网络的接入层而开发的，所以它主要工作在中小型网络的接入层。对网管功能需求不高的环境，可以选择非网管型的S1700系列交换机；对需要像VLAN划分、STP/RSTP、QoS之类的网管功能的环境，则要选择网管型的S1700系列交换机。

1．在企业园区网接入层中的应用

在企业园区网中，S1700 系列交换机可通过百兆/千兆电口（不同机型配置有不同类型的端口，参见表1-8和表1-9）接入终端用户，上行通过千兆光口或百兆/千兆电口 接入汇聚层交换机，进而通过千兆捆绑或万兆上联到骨干网络，构成万兆骨干、百兆到桌面的企业网全网解决方案，满足用户高带宽、多业务的需求。网管型交换机可满足中小型网络对基本交换机功能的配置与管理需求。此种应用的基本网络结构如图1-8所示。

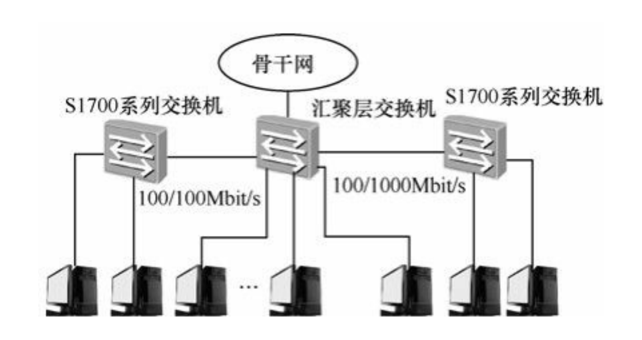


图1-8 S1700系列的企业园区网的接入层应用示例

2．在小型酒店和学校网络中的应用

S1700 系列交换机除了在企业园区网接入层的应用外，还可在一些小型酒店网络和小型学校网络中得到全面应用（可部署在接入层或汇聚层，甚至核心层）。图1-9所示为S1700系列交换机在小型酒店网络中的应用示例，图1-10所示为S1700系列交换机在小型学校网络中的应用示例。

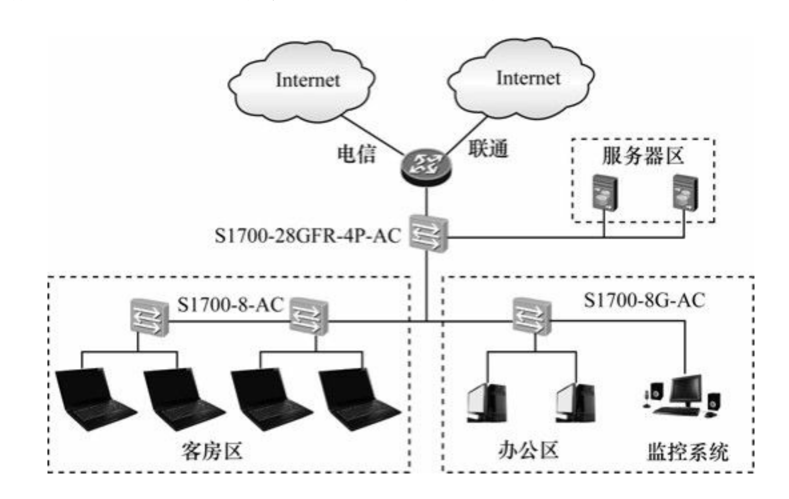


图1-9 S1700系列交换机在小型酒店网络中的应用示例

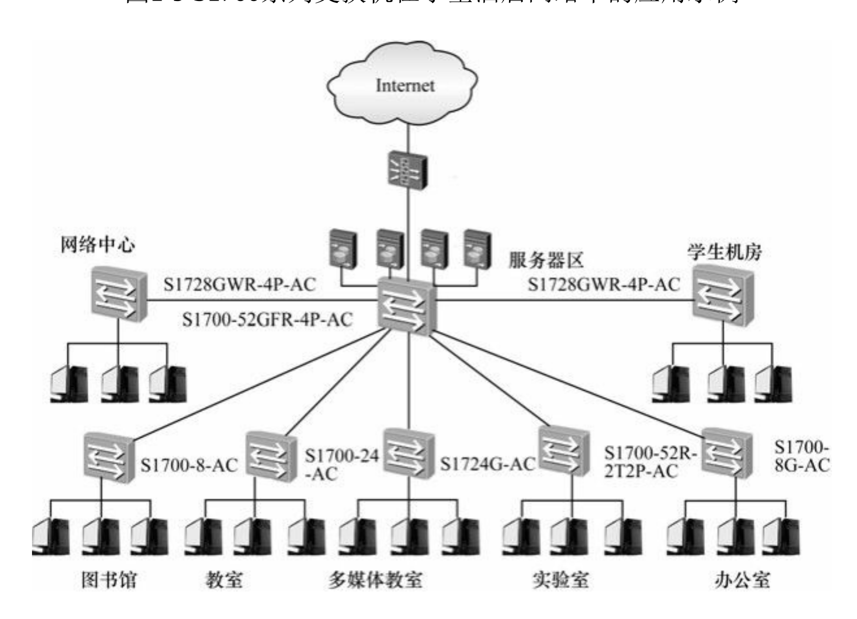


图1-10 S1700系列交换机在小型学校网络中的应用示例

3．在多样化桌面终端接入方面的应用

除上述普通的网络接入功能外，还可利用 S1700网管型交换机提供的Voice VLAN等功能，连接VoIP网关或IP电话，轻松提供多样化的桌面接入功能，如图1-11所示。

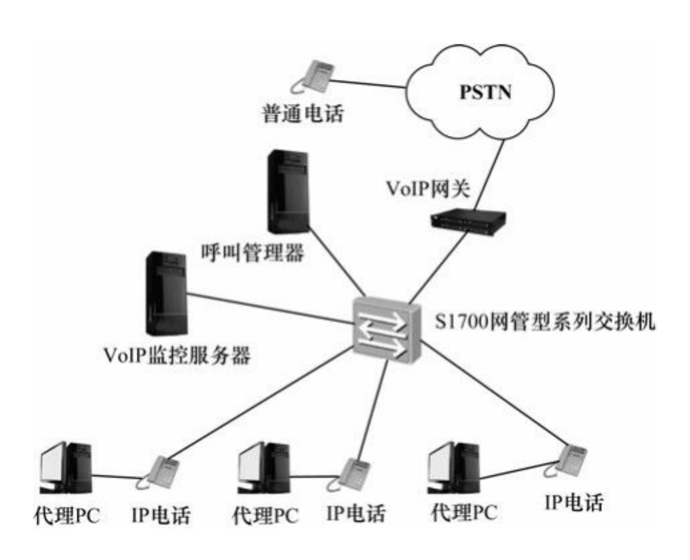


图1-11 S1700网管型系列交换机在多样化终端接入方面的应用示例

1.3 S2700系列交换机的选型与应用

S2700 系列企业交换机是华为公司推出的新一代绿色节能的以太网智能二层百兆（多数有千兆上行端口）接入交换机。它与上节介绍的 S1700 系列交换机一样，也主要定位于企业园区网的接入层。与S1700系列相比，S2700基于华为领先的VRP平台开发，集成了新一代交换技术，具有更强的多业务接入能力、良好的扩展性、丰富的QoS策略、强大的组播复制能力和运营级的安全性，可满足以太网多业务承载和接入需求。

1.3.1 S2700系列机型及基本配置

目前S2700系列主要机型如表1-11所示，分为标准版（SI）和增强版（EI）两种功能版本。相比于S2700-SI系列，S2700-EI系列增加了对QinQ、ACL、802.1x等高级特性的支持。

表1-11 S2700系列机型及基本配置

型号	产品外观	端口数	基本特性
S2700-9TP-SI-AC		9	• 8 个 10/100Base-TX 端口，1 个千兆 Combo 口（10/100/1000Base-T 或 100/1000Base-X，下同） • EI 版本提供交流供电和直流供电两种机型，SI 版本和 PWR 机型提供交流供电，PWR 机型还支持 PoE+ • 转发性能：2.7Mpps • 交换容量：32Gbit/s
S2700-9TP-EI-AC			
S2700-9TP-EI-DC			
S2700-9TP-PWR-EI			

（续表）

型号	产品外观	端口数	基本特性
S2700-18TP-SI-AC S2700-18TP-EI-AC		18	• 16 个 10/100Base-TX 端口，2 个千兆 Combo 口 • 交流供电 • 转发性能：5.4Mpps • 交换容量：32Gbit/s
S2750-20TP-PWR-EI-AC		20	• 16 个 10/100Base-TX 端口，4 个千兆 SFP，2 个复用的千兆 10/100/1000Base-T 以太网端口 Combo 口 • 交流供电，支持 PoE+ • 转发性能：8.4Mpps • 交换容量：32Gbit/s
S2700-26TP-SI-AC S2700-26TP-EI-AC S2700-26TP-EI-DC S2700-26TP-PWR-EI		26	• 24 个 10/100Base-TX 端口，2 个千兆 Combo 口 • EI 版本提供交流供电和直流供电两种机型，SI 版本和 PWR 机型提供交流供电，PWR 机型还支持 PoE+ • 转发性能：6.6Mpps • 交换容量：32Gbit/s
S2750-28TP-EI-AC S2750-28TP-PWR-EI-AC S2751-28TP-PWR-EI-AC		28	• 24 个 10/100Base-TX 以太网端口，4 个千兆 SFP，2 个复用的千兆 10/100/1000Base-T 以太网端口 Combo 口 • 交流供电，PWR 机型还支持 PoE+ • 转发性能：9.6Mpps • 交换容量：32Gbit/s
S2710-52P-SI-AC S2700-52P-EI-AC S2710-52P-PWR-SI S2710-52P-PWR-EI		52	• 48 个 10/100Base-TX 端口，2 个 100/1000Base-X SFP 端口，2 个 1000Base-X SFP 模块上行端口 • 交流供电，PWR 机型还支持 PoE+ • 转发性能：13.2Mpps • 交换容量：32Gbit/s

1.3.2 S2700系列交换机规格及主要特性

S2700系列交换机提供10/100Base-T以太网电口和100/1000Base-X以太网光口，支持Access、Trunk和Hybrid等多种接口类型。对于千兆光纤连接，S2700系列交换机提供了可插拔的SFP类型光模块，光纤长度可以根据用户对传输距离的需求灵活选配。表1-12列出了SI和EI两种功能版本机型的S2700系列机型规格及主要特性。

说明

表中的 S2700-EI*是增强型系列交换机的统称，S2700-SI*是标准型系列交换机的统称，S2710-SI*是S2700-SI的一个子系列。S2700-9TP-SI是S2700-9TP-SI-AC的简写，因产品版本和采用的供电模式无关，因此在描述产品规格时，产品机型名称上没有注明AC或DC，其他型号产品同理。

表1-12 S2700系列交换机的规格及主要特性

<table>
<tr><th colspan="2">项目</th><th>S2700-EI*</th><th>S2700-SI*/S2710-SI*</th></tr>
<tr><td rowspan="2">端口描述</td><td>百兆端口</td><td colspan="2">• S2700-9TP-SI/S2700-9TP-EI/S2700-9TP-PWR-EI：8 个 10/100Base-TX 端口
• S2700-18TP-SI/S2700-18TP-EI：16 个 10/100Base-TX 端口
• S2700-26TP-SI/S2700-26TP-EI/S2700-26TP-PWR-EI：24 个 10/100Base-TX 端口
• S2710-52P-SI/S2700-52P-EI：48 个 10/100Base-TX 端口</td></tr>
<tr><td>千兆端口</td><td colspan="2">• S2700-9TP-SI/S2700-9TP-EI/S2700-9TP-PWR-EI：1 个 GE Combo 端口
• S2710-52P-SI/S2700-52P-EI：2 个 100/1000Base-X，2 个 1000Base-X 端口
• 其他机型：2 个 GE Combo 端口</td></tr>
<tr><td colspan="2">MAC 地址表</td><td colspan="2">• 支持 8k MAC 地址表
• 支持手工添加、删除 MAC 地址表
• 支持 MAC 地址老化时间可配置
• 支持端口/聚合组关闭学习 MAC 能力
• 支持端口 MAC 地址数限制
• 支持黑洞 MAC 地址</td></tr>
<tr><td colspan="2" rowspan="2">VLAN</td><td colspan="2">• 整机支持 4k（1k=1024）个 VLAN
• 支持基于端口的 VLAN</td></tr>
<tr><td>• 支持基于 MAC 地址划分 VLAN
• 支持基于端口的 QinQ</td><td>N/A</td></tr>
<tr><td colspan="2" rowspan="2">QoS</td><td colspan="2">• 支持端口限速和流限速
• 支持每端口 4 个或 8 个优先级队列
• 支持根据报文 802.1p 映射到不同队列
• 支持 SP、WRR、SP+WRR 算法</td></tr>
<tr><td>• 支持基于源 MAC 地址、目的 MAC 地址、源 IP 地址、目的 IP 地址、四层端口、协议类型、VLAN、以太网帧协议、CoS 等信息的 QoS 流分类
• 支持基于流的 802.1p 优先级报文重定向</td><td>N/A</td></tr>
<tr><td colspan="2" rowspan="2">IPv6</td><td colspan="2">• 支持 IPv6 主机功能
• 支持配置静态路由</td></tr>
<tr><td>• 支持 IPv6 ACL
• 支持 MLD v1/v2 Snooping</td><td>N/A</td></tr>
<tr><td colspan="2">组播</td><td colspan="2">• 支持 IGMP v1/v2/v3 Snooping
• 支持捆绑端口的组播负载分担
• 支持基于端口的组播流速率限制和流量统计</td></tr>
<tr><td colspan="2" rowspan="2">端口镜像</td><td colspan="2">支持端口 1:1 或 N:1 镜像</td></tr>
<tr><td>支持基于流镜像</td><td>N/A</td></tr>
<tr><td colspan="2" rowspan="2">安全</td><td>• 支持 802.1x 认证，支持单端口最大用户数限制
• 支持动态 ARP 检测
• 支持 IP Source Guard 功能</td><td>N/A</td></tr>
<tr><td>• 支持 Radius、HWTACACS+、NAC 等多种 AAA 认证方式
• 支持 IP 地址、MAC 地址、交换机端口和 VLAN ID 的组合绑定
• 支持端口限速
• 支持端口隔离、端口安全、Sticky MAC
• 支持包过滤
• 支持 MAC 地址过滤
• 支持多播、广播及未知单播报文抑制
• 支持 MAC 地址学习数目限制
• 支持 CPU 保护功能</td><td>• 支持端口隔离
• 支持多播、广播及未知单播报文抑制
• 支持 CPU 保护功能</td></tr>
</table>

（续表）

<table>
<tr><th>项目</th><th>S2700-EI*</th><th>S2700-SI*/S2710-SI*</th></tr>
<tr><td>管理</td><td colspan="2">• 支持堆叠
• 支持自动配置功能
• 支持 CLI 配置
• 支持 Telnet 远程配置
• 支持 SNMPv1/v2/v3
• 支持 RMON
• 支持集群管理 HGMPv2
• 支持 SSHv2
• 支持 Web 管理特性
• 支持 GVRP 协议</td></tr>
</table>

1.3.3 S2700系列交换机的主要应用

S2700系列交换机与上节介绍的S1700系列交换机一样，都是定位于企业园区网的接入层，但是由于S2700系列交换机集成了华为的VRP系统，所以在功能上更加强大，有着更广泛的应用，可作为大多数企业

园区网的接入层交换机。

1．在企业园区网接入层的应用

图1-12所示为S2700系列交换机作为企业园区网接入层的一个示例。在本示例中， S2700 系列交换机通过百兆电口接入终端用户，上行通过千兆光口或千兆电口接入汇聚层交换机，进而通过千兆捆绑或万兆上联到骨干网络，构成万兆骨干、千兆汇聚、百兆到桌面的企业网全网解决方案，满足用户高带宽、多业务的需求。

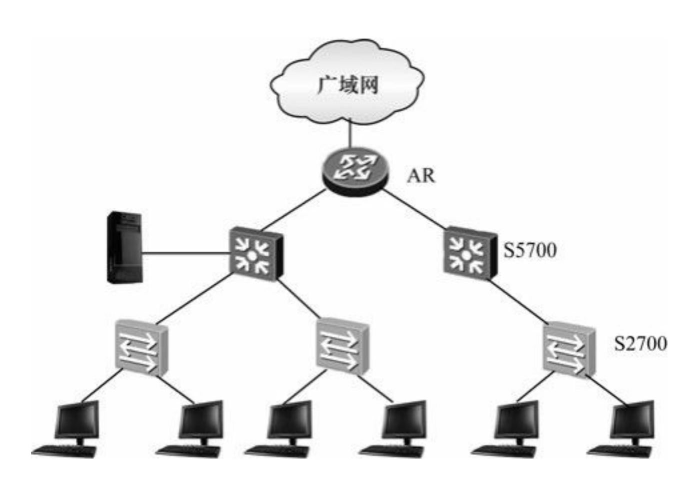


图1-12 S2700系列交换机在企业园区网中的应用示例

2．在多样化桌面终端接入方面的应用

一部分S2700系列机型提供PoE功能，且同时支持802.3af标准和802.3at标准，可为各种大功率的PD设备供电，如 802.11n AP、彩色监控摄像头（Color Camera）、RF读卡器（RF Card Reader）等，可用来构建智能的，无线、有线一体化的企业园区网。

图1-13所示为S2700在多样化桌面终端接入方面的一个应用示例。S2700通过百兆电口接入终端用户，上行通过千兆光口或千兆电口接入汇聚层交换机，进而通过千兆捆绑或万兆上联到骨干网络，构成万兆骨干、百兆到桌面的企业网全网解决方案。在接入 层通过PoE+功能接入各种PD设备。

另外，S2700系列交换机支持基于端口或基于MAC地址划分VLAN（仅EI版本机型支持基于MAC地址划分VLAN）、Voice VLAN（可以接入VoIP设备）；支持Radius、HWTACACS+、NAC（网络访问控制）等多种AAA认证方式；支持IP地址、MAC地址、交换机端口、VLAN ID的组合绑定，支持端口限速；支持端口隔离、端口安全、Sticky MAC，支持包过滤；支持MAC地址过滤等安全功能，可以轻松地提供多样化的桌面接入功能，特别是在移动用户较多、需要部署多种用户访问控制方案的网络环境中。

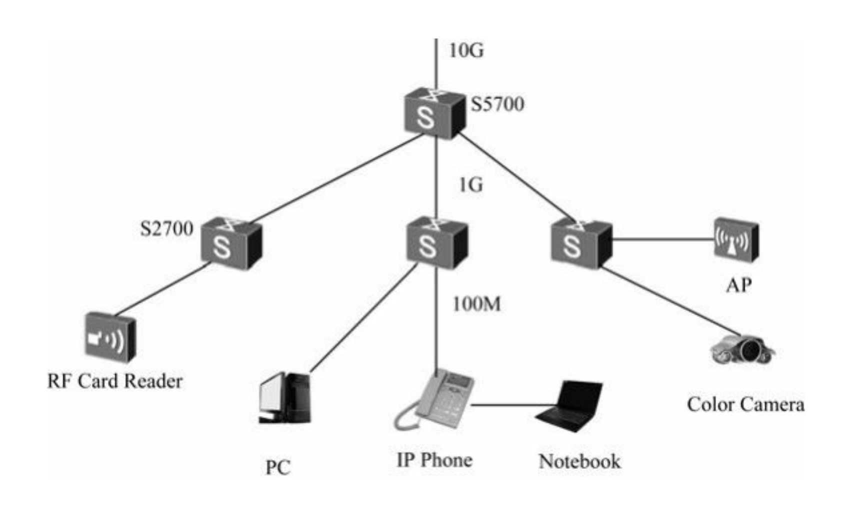


图1-13 S2700系列交换机在多样化桌面终端接入方面的应用示例

1.4 S3700系列交换机的选型与应用

前面介绍的S1700和S2700系列都是属于二层以太网交换机，从S3700系列开始，都是三层以太网交换机了。S3700系列交换机基于新一代高性能硬件和华为 VRP软件平台，主要定位于企业园区网汇聚层或者接入层。它可提供简单便利的安装维护手段、灵活的VLAN部署、PoE供电能力、丰富的路由功能和IPv6平滑升级能力，并通过融合堆叠、虚拟路由器冗余、快速环网保护等先进技术有效增强网络健壮性。

1.4.1 S3700系列机型及基本配置

目前S3700系列的主要机型如表1-13所示，分成标准版（SI）、增强版（EI）和高级版（HI）3种功能版本。

表1-13 S3700系列机型及基本配置

型号	产品外观	端口数	基本配置
S3700-26C-HI		26	• 22 个 10/100Base-TX 端口，2 个千兆 Combo 口（10/100/1000Base-T 或 100/1000Base-X，下同），上行支持 2 个 1000Base-X 插卡 • 双电源，可插拔，可选直流或者交流供电 • 包转发率：9.3Mpps • 交换容量：64Gbit/s

（续表）

型号	产品外观	端口数	基本配置
S3700-28TP-SI-AC		28	• 24 个 10/100Base-TX 端口，2 个 1000Base-X SFP 端口，2 个千兆 Combo 口 • 分交流供电和直流供电两种机型 • 包转发率：9.6Mpps • 交换容量：64Gbit/s
S3700-28TP-EI-AC			
S3700-28TP-SI-DC			
S3700-28TP-EI-DC			
S3700-28TP-EI-MC-AC			• 24 个 10/100Base-TX 端口，2 个 1000Base-X SFP 端口，2 个千兆 Combo 口，2 个监控口 • 交流供电 • 包转发率：9.6Mpps • 交换容量：64Gbit/s
S3700-28TP-PWR-EI			• 24 个 10/100Base-TX 端口，2 个 1000Base-X SFP 端口，2 个千兆 Combo 口 • 双电源，可插拔，交流供电，支持 PoE+ • 包转发率：9.6Mpps • 交换容量：64Gbit/s
S3700-28TP-EI-24S-AC			• 24 个 100Base-FX SFP 端口，2 个 1000Base-X SFP 端口，2 个千兆 Combo 口 • 交流供电 • 包转发率：9.6Mpps • 交换容量：64Gbit/s
S3700-52P-SI-AC		52	• 48 个 10/100Base-TX 端口，2 个 100/1000Base-X SFP 端口，2 个 1000Base-X SFP 端口 • 分交流供电和直流供电两种机型，其中 SI 和 PWR 机型采用交流供电，PWR 机型还支持 PoE+ • 包转发率：13.2Mpps • 交换容量：64Gbit/s
S3700-52P-EI-AC			
S3700-52P-EI-DC			
S3700-52P-PWR-EI			
S3700-52P-EI-48S-AC		52	• 48 个 100Base-FX SFP 端口，2 个 100/1000Base-X SFP 端口，2 个 1000Base-X SFP 端口 • 分交流供电和直流供电两种机型 • 包转发率：13.2Mpps • 交换容量：64Gbit/s
S3700-52P-EI-48S-DC			
S3700-52P-EI-24S-AC			• 24 个 10/100Base-TX 端口，24 个 100Base-FX SFP 端口，2 个 100/1000Base-X SFP 端口，2 个 1000Base-X SFP 端口 • 分交流供电和直流供电两种机型 • 包转发率：13.2Mpps • 交换容量：64Gbit/s
S3700-52P-EI-24S-DC			

1.4.2 S3700系列交换机规格及主要特性

S3700系列交换机中的SI、EI和HI 3个版本的机型规格及主要功能如表 1-14所示。表中的 S3700-SI*是标准型系列交换机的统称，S3700-EI*是增强型系列交换机的统称， S3700-HI*是高级型系列交换机的统称。

表1-14 S3700系列交换机规格及主要功能

<table>
<tr><th colspan="2">项目</th><th>S3700-SI*</th><th>S3700-EI*</th><th>S3700-HI*</th></tr>
<tr><td rowspan="2">端口描述</td><td>百兆端口</td><td colspan="3">• S3700-28TP-EI/S3700-28TP-SI/S3700-28TP-PWR-EI/S3700-28TP-EI-MC：24 个 10/100Base-TX 端口
• S3700-52P-EI/S3700-52P-SI/S3700-52P-PWR-EI：48 个 10/100Base-T 端口
• S3700-28TP-EI-24S：24 个 100Base-FX 端口
• S3700-52P-EI-24S：24 个 10/100Base-T 端口和 24 个 100Base-FX 端口
• S3700-52P-EI-48S：48 个 100Base-FX 端口
• S3700-26C-HI：22 个 10/100Base-T 端口</td></tr>
<tr><td>千兆端口</td><td colspan="3">• SI/EI 28 口设备：2 个 1000Base-X 端口，2 个 GE Combo 端口
• SI/EI 52 口设备：2 个 100/1000Base-X 端口，2 个 1000Base-X 端口
• S3700-26C-HI：2 个 GE Combo 端口</td></tr>
<tr><td colspan="2">扩展插槽</td><td colspan="3">S3700-26C-HI 提供一个扩展插槽，支持上行插卡</td></tr>
<tr><td colspan="2">MAC 地址表</td><td colspan="2">• 支持 16k MAC 地址表
• 支持 MAC 地址自动学习和老化
• 支持静态、动态、黑洞 MAC 表项
• 支持源 MAC 地址过滤</td><td>支持 32k MAC 地址表，其他功能同 EI 版本机型</td></tr>
<tr><td colspan="2">VLAN</td><td colspan="3">• 支持 4k（1k=1024）个 VLAN
• 支持 Guest VLAN、Voice VLAN、Super VLAN
• 支持基于 MAC 地址/IP 协议/IP 子网划分 VLAN
• 支持 QinQ 功能
• 支持灵活 QinQ 功能
• 支持 1:1 和 N:1 VLAN Mapping 功能</td></tr>
<tr><td colspan="2" rowspan="2">可靠性</td><td colspan="3">• 支持 RRPP 环型拓扑、相交环和多实例等功能，且故障保护切换时间低于 50ms
• 支持 SmartLink 树型拓朴及 SmartLink 多实例，提供主备链路的毫秒级保护
• 支持 STP、RSTP 和 MSTP 协议
• S3700-26C-HI 支持 ERPS 以太环保护协议
• 支持 BPDU 保护、根保护和环回保护
• 支持 SEP（Smart Ethernet Protection，智能以太保护）</td></tr>
<tr><td>N/A</td><td colspan="2">支持 BFD For OSPF/ISIS/VRRP/PIM 功能</td></tr>
<tr><td colspan="2" rowspan="2">IP 路由</td><td colspan="3">支持静态路由、RIP v1/v2、ECMP</td></tr>
<tr><td>N/A</td><td colspan="2">OSPF、IS-IS、BGP</td></tr>
<tr><td colspan="2">IPv6</td><td colspan="3">• 支持 ND、PMTU
• 支持 IPv6 Ping、IPv6 Tracert、IPv6 Telnet
• 支持手动配置 Tunnel
• 支持 6to4 tunnel
• 支持 ISATAP tunnel
• 支持基于源 IPv6 地址、目的 IPv6 地址、四层端口、协议类型等 ACL
• 支持 MLD v1/v2 Snooping（Multicast Listener Discovery snooping，组播侦听者发现嗅探）</td></tr>
<tr><td colspan="2" rowspan="2">组播</td><td colspan="2">• 支持 1k 的组播组
• 支持 IGMP v1/v2/v3 Snooping 和快速离开机制
• 支持组播 VLAN 和跨 VLAN 组播复制
• 支持捆绑端口的组播负载分担
• 基于可控组播和基于端口的组播流量统计</td><td>支持 2k 的组播组，其他功能同 EI 版本机型</td></tr>
<tr><td>N/A</td><td colspan="2">支持 IGMP v1/v2/v3、PIM-SM、PIM-DM、PIM-SSM</td></tr>
</table>

（续表）

项目	S3700-SI*	S3700-EI*	S3700-HI*
QoS/ACL	• 支持对端口接收和发送报文的速率进行限制 • 支持报文重定向 • 支持基于端口的流量监管，支持双速三色 CAR 功能，每端口支持 8 个优先级队列 • 支持 WRR、DRR、SP、WRR＋SP、DRR+SP 等队列调度算法 • S3700-26C-HI 机型支持 WRED • 支持报文的 802.1p 和 DSCP 优先级重新标记 • 支持二到四层包过滤功能，提供基于源 MAC 地址、目的 MAC 地址、源 IP 地址、目的 IP 地址、四层端口、协议、VLAN ID 的非法帧过滤功能 • 支持基于队列限速和端口流量整形功能		
安全	• 支持用户分级管理和口令保护 • 支持防止 DoS 攻击、ARP 攻击、ICMP 攻击功能 • 支持 IP 地址、MAC 地址、交换机端口、VLAN ID 的组合绑定 • 支持端口隔离、端口安全、Sticky MAC • 支持黑洞 MAC 地址 • 支持 MAC 地址学习数目限制 • 支持 IEEE 802.1x 认证，支持单端口最大用户数限制 • 支持 Radius、HWTACACS 等多种 AAA 认证方式 • 支持 SSH V2.0 • 支持 CPU 保护功能 • 支持黑名单和白名单		
管理和维护	• 支持智能堆叠（S3700-26C-HI 机型除外） • 支持 MFF • 支持 Telnet 远程配置、维护 • 支持自动配置功能 • 支持 VCT（虚拟电缆检测） • 支持以太网 OAM（运行、管理和维护，即 802.3ah 和 802.1ag 标准） • S3700-28TP-EI-MC-AC 和 S3700-26C-HI 支持断电告警功能 • 支持端口镜像和 RSPAN（远程端口镜像） • 支持 SNMPv1/v2/v3 • 支持 RMON • 支持 MUX VLAN 特性 • 支持 GVRP 协议 • 支持 eSight 网管系统和 Web 管理特性 • 支持自动配置、集群管理 HGMP • 支持 SSH V2 • S3700-26C-HI 支持 HTTPS • S3700-26C-HI 支持 802.3az 能效以太网 EEE		

1.4.3 S3700系列交换机的主要应用

S3700系列交换机是三层交换机系列，它可以应用的领域相对本章前面介绍的S1700和S2700二层交换机系列来说更加广些：既可以作为大型企业园区网的接入层交换设备，又可作为中小型企业园区网的汇聚层，甚至核心层交换设备。

1．在大型企业园区网接入层的应用

S3700系列交换机可以作为大型企业园区网的接入层交换设备。如图1-14所示，在这个大型企业网络中，S3700 系列交换机作为各部门网络的接入层设备，实现百兆桌面接入，然后通过上行的汇聚层交换机S5700系列交换机与核心层交换机S9700系列交换机连接。

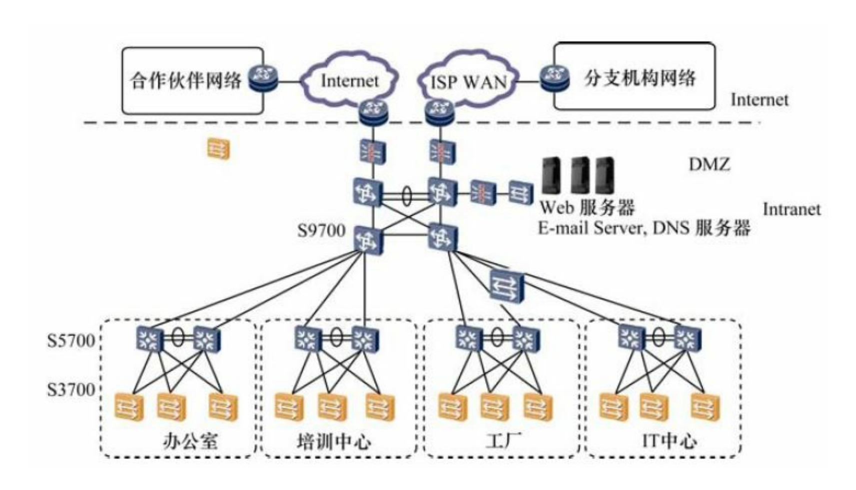


图1-14 S3700系列交换机在大型网络接入层的应用示例

2．在中小型企业园区网核心层的应用

S3700 系列交换机还可用于中小企业网络的核心设备，通过它的路由功能实现跨网段的部门和用户的互访，如图1-15所示。多台S3700系列交换机之间还可以使用堆叠技术实现设备性能和端口的同步扩展。

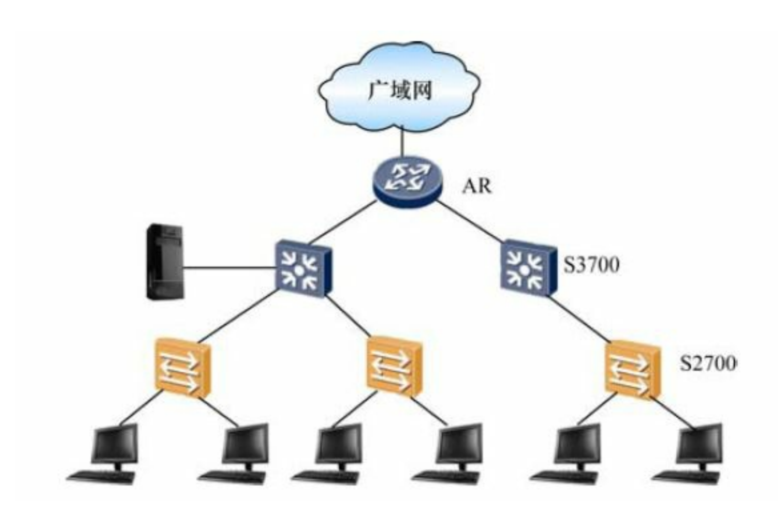


图1-15 S3700系列交换机在中小型网络核心层的应用示例

1.5 S5700系列交换机的选型与应用

S5700 系列交换机是华为公司为满足大带宽接入和以太网多业务汇聚而推出的新一代绿色节能的全千兆以太网交换机。它基于高性能硬件芯片和华为公司领先的VRP软件平台，具备大容量、高密度千兆端口，可提供万兆上行，充分满足企业用户的园区网接入、汇聚、IDC千兆接入以及千兆到桌面等多种应用场景。

1.5.1 S5700系列交换机的机型及基本配置

S5700系列交换机分为LI、SI、EI和HI 4种功能版本，满足于不同用户的需求。这些机型的外观及基本配置如表1-15所示。

表1-15 S5700系列交换机的机型及基本配置

版本	型号	产品外观	端口数	基本配置
L1版本	S5700-10P-LI-AC		10	• 8 个 10/100/1000Base-T 端口，2 个 1000Base-X SFP 端口 • 交流供电，PWR 机型支持 PoE+ • 包转发率：15Mpps • 交换容量：208Gbit/s
	S5700-10P-PWR-LI-AC			
	S5700-28P-LI-AC		28	• 24 个 10/100/1000Base-T 端口，4 个 1000Base-X SFP 端口 • 分交流供电和直流供电两种机型，支持 RPS 冗余电源 • 包转发率：42Mpps • 交换容量：208Gbit/s
	S5700-28P-LI-DC			
	S5700-28X-LI-AC			• 24 个 10/100/1000Base-T 端口，4 个 10GE SFP+端口 • 分交流供电和直流供电两种机型，支持 RPS 冗余电源 • 包转发率：96Mpps • 交换容量：256Gbit/s
	S5700-28X-LI-DC			
	S5700-28TP-LI-AC		28	• 26 个 10/100/1000Base-T 端口，2 个复用的千兆 Combo SFP 口，2 个 1000Base-X 端口 • 交流供电，支持 RPS 冗余电源，PWR 机型还支持 PoE+ • 包转发率：42Mpps • 交换容量：208Gbit/s
	S5700-28TP-PWR-LI-AC			
	S5701-28TP-PWR-LI-AC			
	S5700-28P-PWR-LI-AC			• 24 个 10/100/1000Base-T 端口，4 个 1000Base-X SFP 端口 • 交流供电，支持 RPS 冗余电源和 PoE+ • 包转发率：42Mpps • 交换容量：208Gbit/s
	S5700-28X-PWR-LI-AC			• 24 个 10/100/1000Base-T 端口，4 个 10GE SFP+端口 • 交流供电，支持 RPS 冗余电源和 PoE+ • 包转发率：96Mpps • 交换容量：256Gbit/s
	S5700-52P-LI-AC		52	• 48 个 10/100/1000Base-T 端口，4 个 1000Base-X SFP 端口 • 分交流供电和直流供电两种机型，支持 RPS 冗余电源 • 包转发率：78Mpps • 交换容量：256Gbit/s
	S5700-52P-LI-DC			
	S5700-52X-LI-AC		52	• 48 个 10/100/1000Base-T 端口，4 个 10GE SFP+端口 • 分交流供电和直流供电两种机型，支持 RPS 冗余电源；PWR 机型采用交流供电，支持 RPS 冗余电源和 PoE+ • 包转发率：132Mpps • 交换容量：256Gbit/s
	S5700-52X-LI-DC			
	S5700-52X-PWR-LI-AC			
	S5700-52P-PWR-LI-AC			• 48 个 10/100/1000Base-T 端口，4 个 1000Base-X SFP 端口 • 交流供电，支持 RPS 冗余电源和 PoE+ • 包转发率：78Mpps • 交换容量：256Gbit/s

（续表）

版本	型号	产品外观	端口数	基本配置
L1版本	S5700S-28P-LI-AC		28	• 24 个 10/100/1000Base-T 端口，4 个 1000Base-X SFP 端口 • 交流供电，支持 RPS 冗余电源 • 包转发率：42Mpps • 交换容量：208Gbit/s
	S5700S-52P-LI-AC		52	• 48 个 10/100/1000Base-T 端口，4 个 1000Base-X SFP 端口 • 交流供电，支持 RPS 冗余电源 • 包转发率：78Mpps • 交换容量：256Gbit/s
SI版本	S5700-24TP-SI-AC		24	• 24 个 10/100/1000Base-T 端口，4 个复用的 SFP 千兆端口（Combo） • 分交流供电和直流供电两种机型 • 包转发率：36Mpps • 交换容量：256Gbit/s
	S5700-24TP-SI-DC			
	S5700-24TP-PWR-SI			• 24 个 10/100/1000Base-T 端口，4 个复用的 SFP 千兆端口（Combo） • 可插拔双电源，交流供电，支持 PoE+ • 包转发率：36Mpps • 交换容量：256Gbit/s
	S5700-26X-SI-12S-AC		26	• 12 个 10/100/1000Base-T 端口，12 个 100/1000Base-X 端口，2 个 10GE SFP+口（支持自适应为 GE 光口） • 内置单电源，可外接 RPS • 包转发率：66Mpps • 交换容量：256Gbit/s
	S5700-48TP-SI-AC		48	• 48 个 10/100/1000Base-T 端口，4 个复用的 SFP 千兆端口（Combo） • 分交流供电和直流供电两种机型 • 包转发率：72Mpps • 交换容量：256Gbit/s
	S5700-48TP-SI-DC			
	S5700-48TP-PWR-SI			• 48 个 10/100/1000Base-T 端口，4 个复用的 SFP 千兆端口（Combo） • 可插拔双电源，交流供电，支持 PoE+ • 包转发率：72Mpps • 交换容量：256Gbit/s
	S5700-28C-SI		28	• 24 个 10/100/1000Base-T 端口，4 个复用的 SFP 千兆端口（Combo），上行支持 4×1000Base-X SFP 端口，或 2×10GE SFP+端口，或 4×10GE SFP+插卡 • 可插拔双电源，支持交流或者直流供电；PWR 机型可插拔双电源，交流供电，支持 PoE+ • 包转发率：96Mpps • 交换容量：256Gbit/s
	S5700-28C-PWR-SI			
	S5700-52C-SI		52	• 48 个 10/100/1000Base-T 端口，上行支持 4×1000Base-X SFP 端口，或 2×10GE SFP+端口，或 4×10GE SFP+插卡 • 可插拔双电源，支持交流或者直流供电；PWR 机型可插拔双电源，交流供电，支持 PoE+ • 包转发率：132Mpps • 交换容量：256Gbit/s
	S5700-52C-PWR-SI			

（续表）

版本	型号	产品外观	端口数	基本配置
EI 版本	S5700-28C-EI		28	• 24 个 10/100/1000Base-T 端口，上行支持 4×1000Base-X SFP 端口，或 2×10GE SFP+端口，或 4×10GE SFP+插卡 • 可插拔双电源，支持直流或者交流供电 • 包转发率：96Mpps • 交换容量：256Gbit/s
	S5700-28C-EI-24S			• 24 个 100/1000Base-X 端口，4 个复用的 10/100/1000Base-T 千兆口（Combo），上行支持 4×1000Base-X SFP 端口，或 2×10GE SFP+端口，或 4×10GE SFP+插卡 • 可插拔双电源，支持直流或者交流供电 • 包转发率：96Mpps • 交换容量：256Gbit/s
	S5700-28C-PWR-EI			• 24 个 10/100/1000Base-T 端口，上行支持 4×1000Base-X SFP 端口，或 2×10GE SFP+端口，或 4×10GE SFP+插卡 • 可插拔双电源，交流供电，支持 PoE+ • 包转发率：96Mpps • 交换容量：256Gbit/s
	S5700-52C-EI		52	• 48 个 10/100/1000Base-T 端口，上行支持 4×1000Base-X SFP 端口，或 2×10GE SFP+端口，或 4×10GE SFP+插卡 • 可插拔双电源，支持直流或者交流供电；PWR 机型可插拔双电源，交流供电，支持 PoE+ • 包转发率：132Mpps • 交换容量：256Gbit/s
	S5700-52C-PWR-EI			
	S5710-28C-EI		28	• 24 个 10/100/1000Base-T 端口，4 个复用的 SFP 千兆端口（Combo），或 4 个 10GE SFP+口，上行支持 8×10/100/1000Base-T，或 8×1000Base-X，或 2×10GE SFP+插卡 • 可插拔双电源，支持直流或者交流供电；PWR 机型可插拔双电源，内置 1 个 580W 交流电源，支持 PoE+ • 包转发率：156Mpps • 交换容量：368Gbit/s
	S5710-28C-PWR-EI-AC			
	S5710-52C-EI		52	• 48 个 10/100/1000Base-T 端口，4 个 10GE SFP +口，上行支持 8×10/100/1000Base-T，或 8×1000Base-X，或 2×10GE SFP+插卡 • 可插拔双电源，支持直流或者交流供电；PWR 机型可插拔双电源，其中 S5710-52C-PWR-EI-AC 内置 1 个 580W 交流电源，S5710-52C- PWR-EI 没有内置电源，但都支持 PoE+ • 包转发率：192Mpps • 交换容量：416Gbit/s
	S5710-52C-PWR-EI-AC			
	S5710-52C-PWR-EI			

（续表）

（续表）

版本	型号	产品外观	端口数	基本配置
HI版本	S5700-28C-HI		28	• 24个10/100/1000Base-T端口，上行支持4×1000Base-X SFP端口，或2×10GE SFP+端口，或4×10GE SFP+插卡 • 可插拔双电源，支持直流或者交流供电 • 包转发率：96Mpps • 交换容量：256Gbit/s
	S5700-28C-HI-24S			• 24个100/1000Base-X端口，上行支持4×1000Base-X SFP端口，或2×10GE SFP+端口，或4×10GE SFP+插卡 • 可插拔双电源，支持直流或者交流供电 • 包转发率：96Mpps • 交换容量：256Gbit/s
	S5710-108C-PWR-HI		108	• 48个10/100/1000Base-T，8个10GE SFP+，前面板有3个插槽，支持两种插卡：16×1000Base-X SFP、16×10/100/1000Base-T，后面板有1个插槽，支持以下两种插卡：4×40GE QSFP+、4×10GE SFP+ • 可插拔双电源，交流供电 • 包转发率：504Mpps • 交换容量：1024Gbit/s

1.5.2 S5700系列交换机规格及主要特性

S5700提供精简版（LI）、标准版（SI）、增强版（EI）和高级版（HI）4种产品版本。它们的规格及主要特性如表1-16所示。表中的S5700(S)-LI*是精简型系列交换机的统称， S5700(S)-LI*是 S5700-LI 的一个子系列；S5700-SI*是标准型系列交换机的统称；S5700-EI*是增强型系列交换机的统称，S5710-EI*是S5700-EI的一个子系列；S5700-HI*是高级型系列交换机的统称。

表1-16 S5700系列交换机规格及主要特性

项目		S5700(S)-LI*	S5700-SI*	S5700-EI/S5710-EI*	S5700-HI*
固定端口	LI版本	• S5700-10P-LI-AC/S5700-10P-PWR-LI-AC：8个10/100/1000Base-T端口，2个1000Base-X SFP端口 • S5700-28P-LI/S5700S-28P-LI/S5700-28P-PWR-LI：24个10/100/1000Base-T端口，4个1000Base-X SFP端口 • S5700-28X-LI-AC/S5700-28X-LI-DC/ S5700-28X-PWR-LI-AC：24个10/100/1000Base-T端口，4个10GE SFP＋端口 • S5700-52P-LI/S5700S-52P-LI/S5700-52P-PWR-LI：48个10/100/1000Base-T端口，4个1000Base-X SFP端口 • S5700-52X-LI-AC/S5700-52X-LI-DC/S5700-52X-PWR-LI-AC：48个10/100/1000Base-T端口，4个10GE SFP＋端口			
	SI版本	• S5700-24TP-SI/S5700-24TP-PWR-SI/S5700-28C-SI/S5700-28C-PWR-SI：24个10/100/1000Base-T端口，4个复用SFP千兆端口（Combo） • S5700-26X-SI-12S-AC:12*10/100/1000Base-T，12个100/1000Base-X端口，2个10GE SFP+端口 • S5700-48TP-SI/S5700-48TP-PWR-SI:48个10/100/1000Base-T端口，4个复用SFP千兆端口(Combo) • S5700-52C-SI/S5700-52C-PWR-SI：48个10/100/1000Base-T端口			

（续表）

<table>
<tr><th colspan="2">项目</th><th>S5700(S)-LI*</th><th>S5700-SI*</th><th>S5700-EI/S5710-EI*</th><th>S5700-HI*</th></tr>
<tr><td rowspan="2">固定端口</td><td>EI版本</td><td colspan="4">• S5700-28C-EI/S5700-28C-PWR-EI：24 个 10/100/1000Base-T 端口
• S5700-52C-EI/S5700-52C-PWR-EI：48 个 10/100/1000Base-T 端口
• S5700-28C-EI-24S：24 个 100/1000Base-X 端口，4 复用 10/100/1000Base-T(Combo) 端口
• S5710-28C-EI/ S5710-28C-PWR-EI-AC：24 个 10/100/1000Base-T 端口，4 个复用的 SFP 千兆端口（Combo），4 个 10GE SFP +端口
• S5710-52C-EI/ S5710-52C-PWR-EI-AC/S5710-52C-PWR-EI：48 个 10/100/1000Base-T 端口，4 个 10GE SFP +端口</td></tr>
<tr><td>HI版本</td><td colspan="4">• S5700-28C-HI：24 个 10/100/1000Base-T 端口
• S5700-28C-HI-24S：24 个 100/1000Base-X 端口</td></tr>
<tr><td colspan="2">扩展插槽</td><td colspan="4">• S5700TP 系列机型提供一个堆叠扩展插槽
• S5700C 系列机型提供两个扩展插槽，分别支持上行插卡和堆叠卡
• S5710C 系列机型提供两个扩展插槽，支持上行插卡
• S5700HI 系列机型提供一个扩展插槽，支持上行插卡</td></tr>
<tr><td colspan="2" rowspan="2">MAC 地址表</td><td colspan="2">支持 16k MAC 地址容量(但 S5700S-LI 仅支持 8k MAC 地址容量）</td><td colspan="2">支持 32k MAC 地址容量</td></tr>
<tr><td colspan="4">• 支持 MAC 地址自动学习和老化
• 支持静态、动态、黑洞 MAC 表项
• 支持源 MAC 地址过滤</td></tr>
<tr><td colspan="2">VLAN 特性</td><td colspan="4">• 支持 4k（1k=1024）个 VLAN
• 支持 Guest VLAN、Voice VLAN
• 支持基于 MAC/协议/IP 子网/策略/端口划分 VLAN
• 支持 1:1 和 N:1 VLAN 映射功能
• 支持 SuperVLAN（S5700(S)-LI 除外）</td></tr>
<tr><td colspan="2" rowspan="2">可靠性</td><td colspan="4">• 支持 RRPP 环型拓扑和 RRPP 多实例
• 支持 SmartLink 树型拓朴和 SmartLink 多实例，提供主备链路的毫秒级保护
• 支持智能以太保护 SEP 协议
• 支持 STP/RSTP/MSTP 协议
• 支持 BPDU 保护、根保护和环回保护
• 支持 Enhanced Trunk（S5700(S)-LI 系列除外）</td></tr>
<tr><td colspan="2">NA</td><td colspan="2">支持 BFD For OSPF/ISIS/VRRP/PIM 协议</td></tr>
<tr><td colspan="2">MPLS 特性</td><td colspan="2">NA</td><td colspan="2">• 支持 MPLS L3VPN
• 支持 MPLS L2VPN（VPWS/VPLS）
• 支持 MPLS-TE
• 支持 MPLS QoS</td></tr>
<tr><td colspan="2">IP 路由特性</td><td>支持静态路由</td><td>支持静态路由、RIP v1/2、RIPng、ECMP、路由策略</td><td colspan="2">支持静态路由、RIP v1/2、RIPng、OSPF、OSPFv3、IS-IS、IS-ISv6、BGP、BGP4+、ECMP、路由策略</td></tr>
<tr><td colspan="2">IPv6 特性</td><td colspan="4">• 支持 ND、PMTU
• 支持 IPv6 Ping、IPv6 Tracert、IPv6 Telnet
• 支持基于源 IPv6 地址、目的 IPv6 地址、四层端口、协议类型等 ACL
• 支持 MLD v1/v2 snooping（Multicast Listener Discovery snooping）
• 支持 6to4、ISATAP、手动配置 tunnel（S5700(S)-LI 除外）</td></tr>
<tr><td colspan="2" rowspan="2">组播特性</td><td colspan="4">• 支持 IGMP v1/v2/v3 Snooping 和快速离开机制
• 支持 VLAN 内组播转发和组播多 VLAN 复制
• 支持捆绑端口的组播负载分担
• 支持可控组播
• 支持基于端口的组播流量统计</td></tr>
<tr><td colspan="2">NA</td><td colspan="2">支持 IGMP v1/v2/v3、PIM-SM、PIM-DM、PIM-SSM 和 MSDP</td></tr>
</table>

（续表）

项目	S5700(S)-LI*	S5700-SI*	S5700-EI/S5710-EI*	S5700-HI*
QoS/ACL特性	• 支持对端口接收和发送报文的速率进行限制 • 支持报文重定向 • 支持基于端口的流量监管，支持双速三色 CAR 功能 • 每端口支持 8 个队列 • 支持 WRR、DRR、SP、WRR＋SP、DRR+SP 队列调度算法 • S5710-EI 和 S5700-HI 系列机型还支持 WRED（Weighted Random Early Detection，加权随机先期检测） • 支持报文的 802.1p 和 DSCP 优先级重标记 • 支持二到四层的包过滤功能，提供基于源 MAC 地址、目的 MAC 地址、源 IP 地址、目的 IP 地址、端口、协议、VLAN 的非法帧过滤功能 • 支持基于队列限速和端口流量整形功能			
安全特性	• 支持用户分级管理和口令保护 • 支持防止 DoS、ARP 攻击功能、ICMP 防攻击 • 支持 IP 地址、MAC 地址、端口、VLAN ID 的组合绑定 • 支持端口隔离、端口安全、Sticky MAC • 支持黑洞 MAC 地址 • 支持 MAC 地址学习数目限制 • 支持 IEEE 802.1x 认证和单端口最大用户数限制 • 支持 Radius、HWTACACS、NAC 等多种 AAA 认证方式 • 支持 SSH V2.0 • 支持 HTTPS • 支持 CPU 保护功能 • 支持黑名单和白名单			
OAM	支持			• 采用硬件实现，支持 EFM OAM、CFM OAM • Y.1731 性能检测：支持硬件级时延和抖动检测
管理和维护	• 支持智能堆叠（S5700-HI 和 S5700S-LI 系列机型除外） • 支持 MFF 和虚拟电缆检测（Virtual Cable Test） • 支持端口镜像和 RSPAN（远程端口镜像） • 支持 Telnet 远程配置、维护 • 支持 SNMPv1/v2/v3 • 支持 RMON • 支持 eSight 网管系统、支持 Web 网管特性 • 支持自动配置 • 支持集群管理 HGMP • 支持 HTTPS • 支持系统日志、分级告警 • 支持 GVRP 协议 • 支持 MUX VLAN 功能 • 支持 802.3az 能效以太网 EEE（S5700-LI 和 S5700-HI 支持） • 支持断电告警 Dying gasp 功能（S5700-LI 支持） • 支持 NetStream（S5710-EI/S5700-HI 支持）			
	NA		支持 sFlow	

1.5.3 S5700系列交换机的主要应用

由于S5700系列交换机的功能和性能非常强大，因此它可以在更大型的网络和更高级的工作层次上得到应用。如在大型企业园区网或数据中心中作为千兆接入层交换机，在中小型企业园区网中作为核心层或者汇聚层交换机。

1．在大型企业园区网汇聚层中的应用

S5700系列交换机在大型企业园区网汇聚层中的应用参见1.4.4节的图1-14。

2．在数据中心接入层中的应用

S5700系列交换机在中型数据中心接入层中的应用如图1-16所示。通过它可接入千兆服务器，再利用链路捆绑汇聚到上层设备。单机架服务器超过一定数量可采用 S5700系列交换机的iStack堆叠功能完成服务器接入，便于维护和提高可靠性。

3．在大型企业园区网接入层的应用

在大型企业园区网接入层中，S5700 系列交换机可实现千兆到桌面的接入，满足高性能终端用户的接入需求，如图1-17所示。

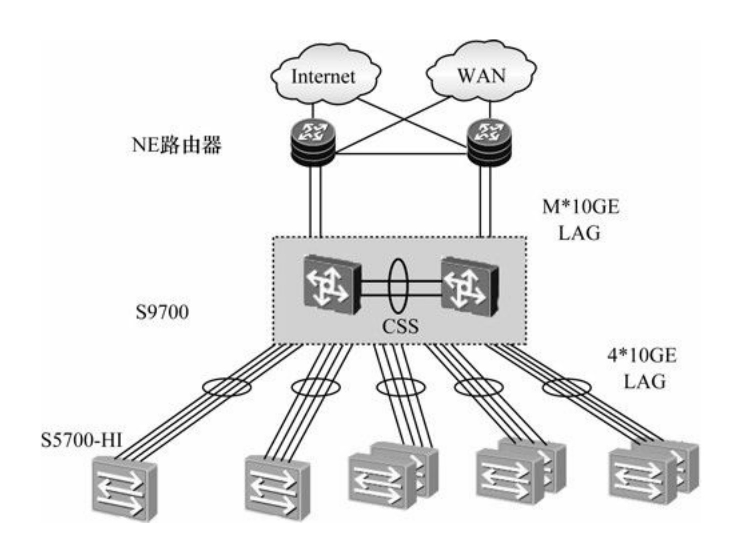


图1-16 S5700系列交换机在数据中心接入层中的应用示例

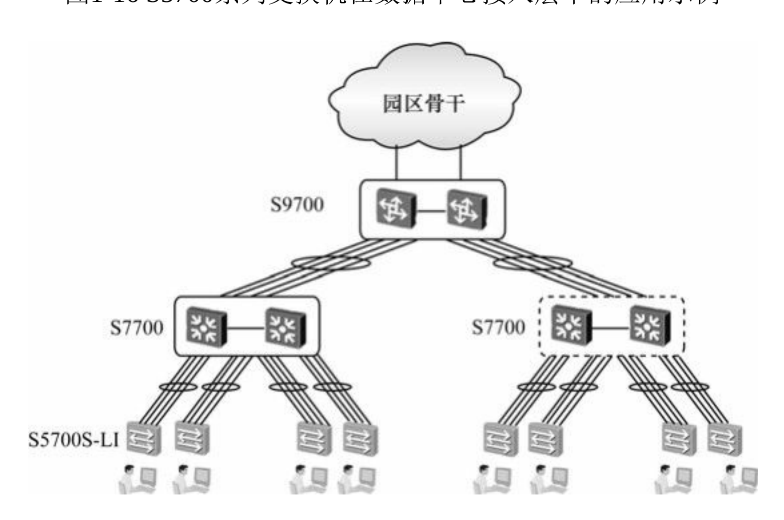


图1-17 S5700系列交换机在大型企业园区网接入层中的应用

1.6 S6700系列交换机的选型与应用

S6700 系列交换机是华为公司最新开发的下一代全万兆盒式交换机，可用于数据中 心万兆服务器接入及园区网的核心。它是业内最高性能的盒式交换机之一，同时提供最多 24/48 个全线速万兆接口，使万兆服务器高密度接入和园区网高密度万兆汇聚成为可能。同时，S6700支持丰富的业务特性、完善的安全控制策略、丰富的QoS等特性以满足数据中心可扩展性、可靠性、可管理性、安全性等诸多挑战。

1.6.1 S6700系列机型及基本配置

S6700 系列交换机主要应用于数据中心万兆接入和企业园区网汇聚层或核心层。目前只有S6700-24-EI和S6700-48-EI两款机型，其外观及基本配置如表1-17所示。

表1-17 S6700系列机型及基本配置

产品外观	基本配置
S6700-48-EI	• 48 个 GE SFP/10 GE SFP+端口 • 可插拔双电源，支持交流或者直流供电 • 支持 USB • 包转发率：715Mpps • 交换容量：960Gbit/s
S6700-24-EI	• 24 个 GE SFP/10 GE SFP+端口 • 可插拔双电源，支持交流或者直流供电 • 支持 USB • 包转发率：358Mpps • 交换容量：480Gbit/s

1.6.2 S6700系列交换机的规格及主要特性

S6700系列交换机的两款机型规格及基本功能如表1-18所示。

表1-18 S6700系列交换机的规格及主要功能

项目	S6700-24-EI	S6700-48-EI
端口描述	24 个 GE SFP/10 GE SFP+端口	48 个 GE SFP/10 GE SFP+端口
MAC 地址表	• 最大支持 128k（1k=1024）个 MAC 地址 • 支持 MAC 地址自动学习和老化 • 支持静态、动态、黑洞 MAC 表项 • 支持源 MAC 地址过滤	
VLAN 特性	• 支持 4k（1k=1024）个 VLAN • 支持 Guest VLAN、Voice VLAN • 支持基于 MAC/协议/IP 子网/策略/端口的 VLAN • 支持 1:1 和 *N*:1 VLAN 交换功能 • 支持基本、灵活 QinQ 功能	
IPv4 路由	• 支持静态路由，以及 RIP v1/2、OSPF、IS-IS、BGP 动态路由 • 支持 ECMP（等价开销多路径）和 URPF（单播反向路径转发） • 支持 VRRP（虚拟路由器冗余协议） • 支持策略路由 • 支持路由策略	
IPv6 路由	• 支持静态路由和 OSPFv3、BGP4+ • 支持 RIPng	
IPv6 特性	• 支持 ND（Neighbor Discovery） • 支持 PMTU	

（续表）

（续表）

项目	S6700-24-EI	S6700-48-EI
IPv6 特性	• 支持 IPv6 Ping、IPv6 Tracert、IPv6 Telnet • 支持 6to4、ISATAP、手动配置隧道 • 支持基于源 IPv6 地址、目的 IPv6 地址、四层端口、协议类型等 ACL • 支持 MLD v1/v2 snooping	
组播	• 支持二层静态组播 MAC • 支持 MAC 模式转发 • 支持 IGMP Snooping 和快速离开机制 • 支持组播 VLAN • 支持 MLD Snooping • 支持 IGMP Proxy • 支持可控组播 • 基于端口的组播流量统计 • 支持 IGMP v1/v2/v3 • 支持 PIM-SM、PIM-DM、PIM-SSM • 支持 MSDP	
QoS/ACL	• 支持对端口接收和发送报文的速率进行限制 • 支持报文重定向 • 支持基于端口的流量监管，支持双速三色 CAR 功能 • 每端口支持 8 个队列 • 支持 WRR、DRR、SP、WRR＋SP、DRR+SP 队列调度算法 • 支持 WRED • 支持报文的 802.1p 和 DSCP 优先级重新标记 • 支持二到四层的包过滤功能，提供基于源 MAC 地址、目的 MAC 地址、源 IP 地址、目的 IP 地址、端口、协议、VLAN 的非法帧过滤功能 • 支持基于队列限速和端口整形功能	
可靠性	• 支持 STP、RSTP 和 MSTP 协议 • 支持 BPDU 保护、根保护和环回保护 • 支持 RRPP 环型拓扑和 RRPP 多实例 • 支持 SmartLink 树型拓扑和 SmartLink 多实例，提供主备链路的毫秒级保护 • 支持智能以太保护协议（SEP） • 支持 ERPS 以太环保护协议（G.8032） • 支持 BFD for OSPF/ISIS/VRRP/PIM 协议 • 支持增强 Trunk（E-trunk）	
安全特性	• 支持用户分级管理和口令保护 • 支持防止 DoS、ARP 攻击功能、ICMP 防攻击 • 支持 IP 地址、MAC 地址、连接端口、VLAN ID 的组合绑定 • 支持端口隔离、端口安全、Sticky MAC • 支持黑洞 MAC 地址 • 支持 MAC 地址学习数目限制 • 支持 IEEE 802.1x 认证，支持单端口最大用户数限制 • 支持 AAA 认证，支持 Radius、HWTACACS、NAC 等多种方式 • 支持 SSH V2.0 • 支持 HTTPS • 支持 CPU 保护功能 • 支持黑名单和白名单	
管理和维护	• 支持堆叠（业务口实现） • 支持 MAC 地址强制转发（MFF） • 支持虚拟电缆检测（VCT）	

（续表）

项目	S6700-24-EI	S6700-48-EI
管理和维护	• 支持以太网 OAM（802.3ah 和 802.1ag） • 支持本地端口镜像和远程端口镜像（RSPAN），支持观察端口正常转发报文 • 支持 Telnet 远程配置、维护 • 支持 SNMPv1/v2/v3 • 支持 RMON • 支持网管系统、支持 Web 网管特性 • 支持集群管理 HGMP • 支持系统日志、分级告警 • 支持 GVRP 协议 • 支持 MUX VLAN 功能 • 支持 sFlow	

1.6.3 S6700系列交换机的应用

S6700系列交换机主要应用于大型数据中心和大中型企业园区网的汇聚层或核心层。

1．在大型数据中心中的应用

S6700系列交换机在大型数据中心中的应用如图1-18所示。此时采用T比特路由交换平台 S9700 系列交换机作为数据中心核心，通过集成的防火墙、负载均衡等多业务板卡来实现安全保证和流量均衡。万兆服务器接入可使用 S6700 提供高密度的万兆接入方案。

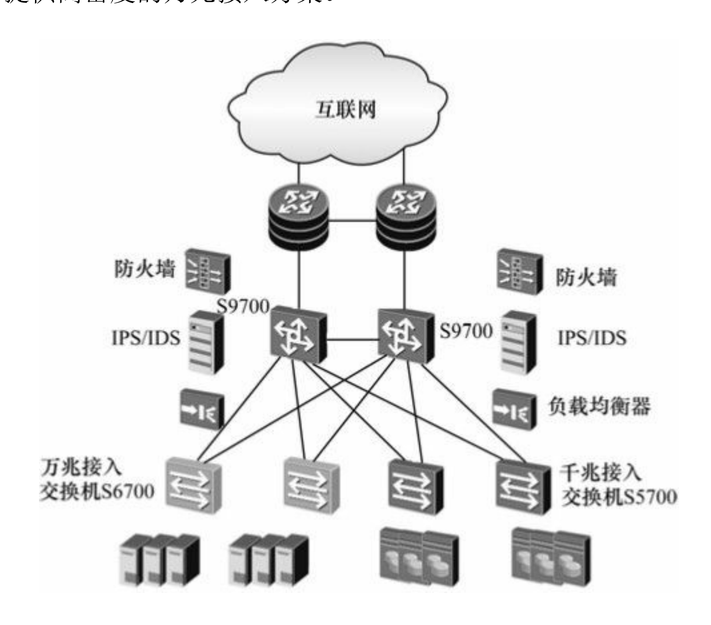


图1-18 S6700系列交换机在数据中心中的应用示例

2．在大型企业园区网中的应用

S6700系列交换机可用于大型企业园区网的汇聚层或核心层，如图1-19所示。其业界领先的高密度全线速万兆端口，为上、下行高速连接提供了完善的解决方案；其丰富的业务特性、完善的安全控制机制使得S6700系列交换机成为园区交换机最高性价比的选择。

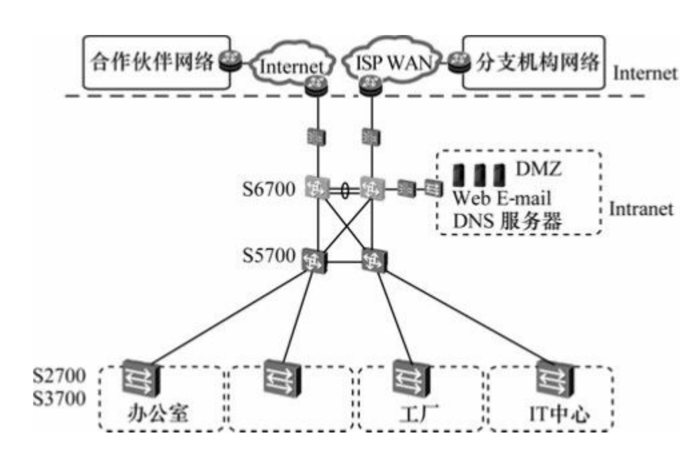


图1-19 S6700系列交换机在大型网络中的应用示例

1.7 S7700/9300/9700系列交换机的选型与应用

S7700、S9300和S9700三大系列都是华为公司面向下一代企业网络架构而推出的新一代高端智能T比特

路由交换机。这三大系列交换机产品基于华为公司智能多层交换的技术理念，在提供稳定、可靠、安全的高性能二到四层交换服务基础上，进一步融合了MPLS VPN、硬件 IPv6、桌面云、视频会议、无线等多种网络业务，提供不间断升级、不间断转发、硬件 OAM/BFD、环网保护等多种高可靠技术，广泛适用于园区网络和数据中心网络，可对无线、话音、视频和数据融合网络进行先进的控制，帮助企业构建交换路由一体化的端到端融合网络。

因为这三大系列的主要功能特性差不多，主要是性能上的差异，故在此一起介绍并进行横向比较，可更方便用户的选择。

1.7.1 S7700/9300/9700系列交换机规格

S7700、S9300和S9700三大系列都是机箱式结构，都提供3、6和12三种不同业务槽机型选择，它们的配置规格比较如表1-19所示。

表1-19 S7700、S9300和S9700系列的三种机型配置规格比较

项目	S7703	S9303	S9703	S7706	S9306	S9706	S7712	S9312	S9712
背板容量	3Tbit/s		7.2 Tbit/s	6Tbit/s		14.4 Tbit/s	12Tbit/s		19.2 Tbit/s
交换容量	1.92 Tbit/s	768 Gbit/s/ 1.92 Tbit/s	2.88 Tbit/s/ 5.76 Tbit/s	3.84 Tbit/s/ 5.12 Tbit/s	2 Tbit/s/ 5.12 Tbit/s	6.72 Tbit/s/ 14.72 Tbit/s	3.84 Tbit/s/ 5.76 Tbit/s	2 Tbit/s/ 5.12 Tbit/s	8.64 Tbit/s/ 18.56 Tbit/s
包转发率	576Mpps/ 1440Mpps		2160 Mpps	1152Mpps/ 2880Mpps		2880Mpps/ 5040Mpps	1344Mpps/ 3360Mpps		3840Mpps/ 6480Mpps

1.7.2 S7700/9700系列交换机的主要特性

S7700、S9300和S9700三大交换机系列的主要特性也基本一样（主要是硬件配置与 培训中心
性能上的差异），具体如表1-20所示。

表1-20 S7700、S9300和S9700三大系列主要功能特性

项目	S7700	S9300/S9700
VLAN	• 支持 Access、Trunk、Hybrid 端口类型 • 支持 default VLAN • 支持 VLAN 交换 • 支持 QinQ、增强型灵活 QinQ • 支持基于 MAC 的动态 VLAN 分配	
MAC 地址功能	• 支持 MAC 地址自动学习和老化 • 支持静态、动态、黑洞 MAC 表项 • 支持源 MAC 地址过滤 • 支持基于端口和 VLAN 的 MAC 地址学习限制	
STP	• 支持 STP(IEEE 802.1d)，RSTP(IEEE 802.1w)和 MSTP(IEEE 802.1s) • 支持 BPDU 保护、Root 保护、环路保护 • 支持 BDPU Tunnel	
IP 路由	• 支持 RIP、OSPF、ISIS、BGP 等 IPv4 动态路由协议 • 支持 RIPng、OSPFv3、ISISv6、BGP4+等 IPv6 动态路由协议	
组播	• 支持 IGMPv1/v2/v3、IGMP v1/v2/v3 Snooping • 支持 PIM DM、PIM SM、PIM SSM • 支持 MSDP、MBGP • 支持用户快速离开机制 • 支持组播流量控制 • 支持组播查询器 • 支持组播协议报文抑制功能 • 支持组播 CAC • 支持组播 ACL	
MPLS	• 支持 MPLS 基本功能 • 支持 MPLS OAM • 支持 MPLS TE • 支持 MPLS VPN/VLL/VPLS	
可靠性	• 支持 LACP、支持跨设备 E-Trunk • 支持 VRRP、BFD for VRRP • 支持 BFD for BGP/IS-IS/OSPF/静态路由 • 支持 NSF、GR for BGP/IS-IS/OSPF/LDP • 支持 TE FRR、IP FRR • 支持以太网 OAM 802.3ah 和 802.1ag • 支持 ITU-Y.1731 • 支持 DLDP • 支持运行中软件升级 ISSU • 支持集群交换系统 CSS	
QoS	• 支持基于 Layer2 协议头、Layer3 协议、Layer4 协议、802.1p 优先级等的组合流分类 • 支持 ACL、CAR、Remark、Schedule 等动作 • 支持 PQ、WRR、DRR、PQ+WRR、PQ+DRR 等队列调度方式 • 支持 WRED、尾丢弃等拥塞避免机制 • 支持流量整形	
	N/A	支持 H-QoS
配置与维护	• 支持 Console、Telnet、SSH 等终端服务 • 支持 SNMPv1/v2/v3 等网络管理协议 • 支持通过 FTP、TFTP 方式上载、下载文件 • 支持 BootROM 升级和远程在线升级 • 支持热补丁 • 支持用户操作日志	

（续表）

项目	S7700	S9300/S9700
安全和管理	• 支持 802.1x 认证和 Portal 认证 • 支持 NAC • 支持 RADIUS 和 HWTACACS 用户登录认证 • 支持命令行分级保护，未授权用户无法侵入 • 支持防范 DoS 攻击、TCP 的 SYN Flood 攻击、UDP Flood 攻击、广播风暴攻击、大流量攻击 • 支持 1kB CPU 通道队列保护 • 支持 ICMP 实现 ping 和 traceroute 功能 • 支持 RMON	
增值业务能力	• 支持 Firewall 功能 • 支持 NAT 功能 • 支持 Netstream 功能 • 支持 IPSec 功能 • 支持负载均衡功能 • 支持无线 AC 控制器	

1.7.3 S7700系列交换机的应用

S7700系列交换机主要应用于大型企业园区网的汇聚层和中小型企业园区网核心层。

1．在大型企业园区网汇聚层中的应用

S7700系列智能路由交换机具备5.12Tbit/s交换容量和高密度万兆端口，可以作为大型企业园区汇聚交换机设备，组建高可靠、业务易扩展、易管理的企业园区网络，如图1-20所示。S7700支持硬件CPU通道队列，防火墙功能模块，在汇聚层为企业业务增加安全保障，保护企业核心免受DDoS攻击与各种安全威胁。

2．在中小型企业园区网核心层中的应用

S7700系列智能路由交换机支持OSPF、BGP等路由协议，支持IP、MPLS全线速转发，同时具备防火墙、IPSec模块，整机5.12Tbit/s交换容量，可以作为中小型企业园区网络的核心设备，如图1-21所示。它可为中小企业园区网提供高性价比、高可靠、多业务易部署的网络解决方案。

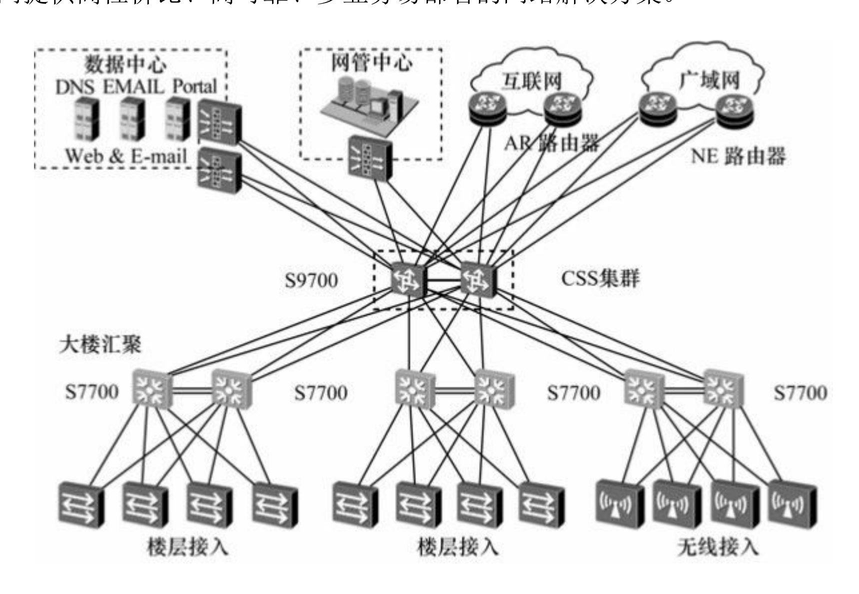


图1-20 S7700系列交换机在大型网络汇聚层中的应用示例

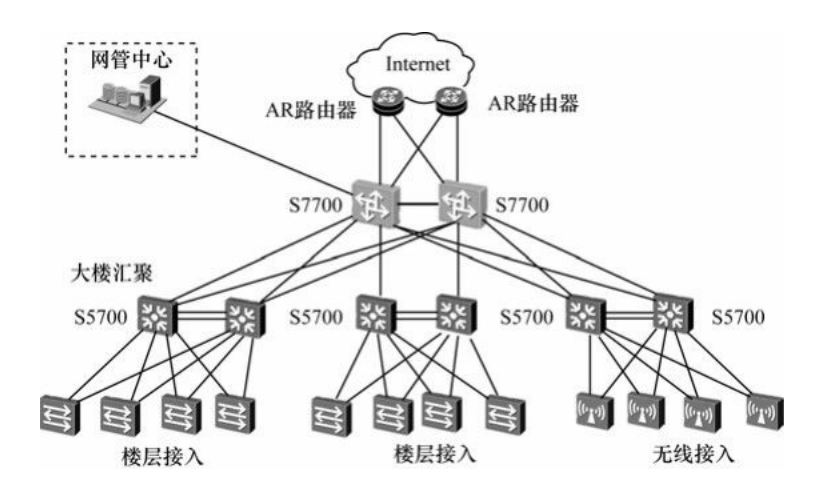


图1-21 S7700系列交换机在中小型网络核心层中的应用示例

1.7.4 S9300/9700系列交换机的主要应用

S9300和S9700两大系列交换机主要在大型企业园区网的核心层，以及大型数据中心核心层或汇聚层中应用。下面分别予以介绍。

1．在大型企业园区网核心层中的应用

S9300和S9700两大系列交换机主要作为大型企业园区网的核心层交换设备，能够为用户组建高可靠、高性能、业务易扩展、易管理的企业园区网，如图1-22所示。

在这种应用中，主要利用了它们以下的功能特性。

（1）具备分布式IPv4/IPv6/MPLS全线速交换能力，满足企业园区核心、汇聚节点高密万兆海量数据吞吐能力。

（2）利用CSS集群、负载均衡一体化解决方案，可大大提高网络IT利用效率，降低网络维护成本。

（3）利用支持的无线AC控制模块，又可使园区核心与WLAN控制合一，节省建网投资。

（4）利用支持的硬件CPU通道队列，可保护企业核心免受DDoS攻击与各种安全威胁。

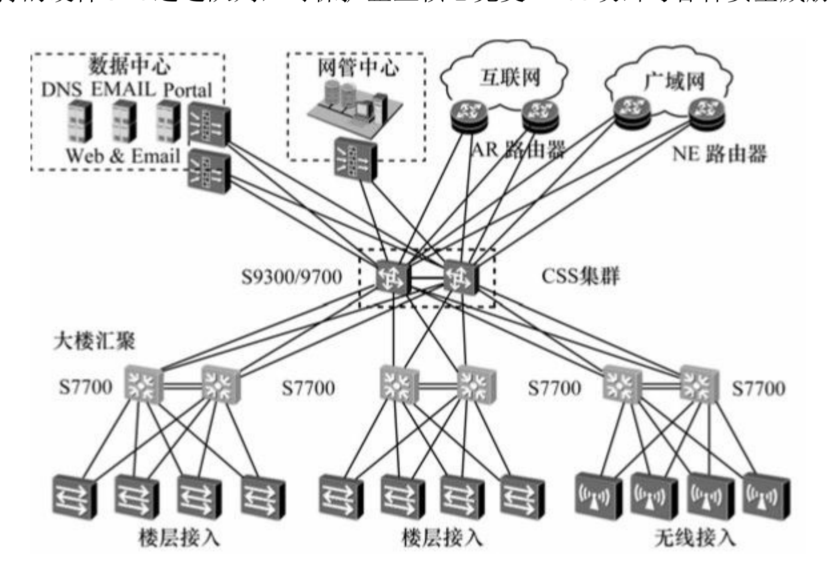


图1-22 S9300/9700系列交换机在大型网络核心层中的应用示例

2．在大型数据中心核心层的应用

S9300和S9700两大系列交换机还可以作为大型数据中心的高密万兆核心和万兆汇聚节点，助力企业构筑高可靠、无阻塞、虚拟化的数据中心网络，如图1-23所示。利用它们所支持的 ISSU、IP FRR、硬件级BFD、NSF、VRRP、E-Trunk等高可靠性技术，可实现数据中心业务永续运行。同样利用它们所提供的CSS集群、负载均衡一体化解决方案，可提高网络IT利用效率。

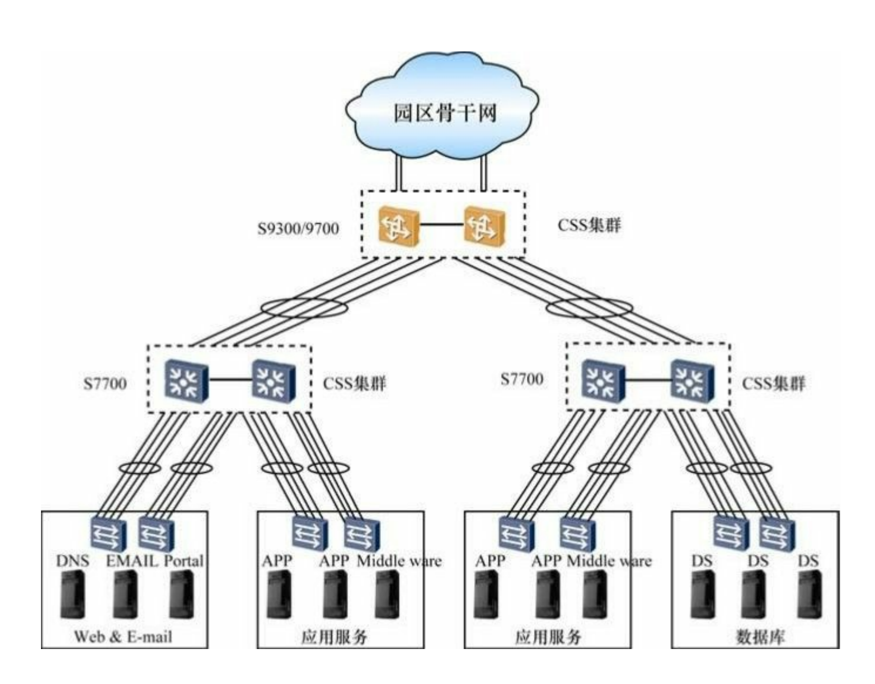


图1-23 S9300/9700系列交换机在大型数据中心核心层中的应用示例

第2章　VRP系统基础及基本使用

对于一个新认识的网络交换机，首先要学习的就是该交换机的网络操作系统本身，了解它的基本使用和管理方法，否则后面的许多配置与管理方法学习将无从谈起。

本章作为正式介绍华为S系列园区交换机的开始，首先介绍S系列园区交换机操作系统——VRP系统（以当前Sx700系列所用的VRP 5.x版本为例）本身的一些基础知识，如VRP命令行格式、命令行视图、命令行基本使用与操作，以及VRP文件系统管理、VRP系统软件/配置文件管理、交换机启动设置等。虽然本章内容很基础，但对于一个初学者来说仍然非常重要。

2.1 VRP系统基础

VRP（Versatile Routing Platform，通用路由平台）是华为公司数据通信产品的通用网络操作系统平台，包括路由器、交换机、防火墙、WLAN等众多系列产品。

2.1.1 VRP系统概述

VRP系统自1994年开始研发至今已走过了近20年的历史，系统版本也从最初的1.x发展到了现在最高的8.x 版本，无论是从系统软件体系结构，还是从支持的功能，采用的配置方法上发生了明显的变化。在S系列以太网交换机中目前主要是应用5.x版本， VRP 8.x目前主要应用在数据交换机CE系列和集群路由器NE5000E。VRP系统可以运行在多种硬件平台之上并拥有一致的网络界面、用户界面和管理界面，为用户提供了灵活丰富的应用解决方案。

当然VRP系统的发展远不仅体现在其版本、功能上的更新，更体现在软件平台结构上的发展、变化，从VRP 1.x的集中式到VRP 3.x和今天主流应用的VRP 5.x的模块化分布式，在平台架构上不断优化的同时也大大提升了平台的性能，降低了产品成本。下一代的VRP 8.x还采用了多进程、多处理、内存保护等新技术。

VRP以TCP/IP协议栈为核心，在操作系统中集成了路由、组播、QoS、VPN、安全和IP话音等数据通信要件。VRP平台是以当前最主流的IP业务为核心，实现组件化的体系结构。在提供丰富功能特性的同时，还提供基于应用的可裁剪能力和可伸缩能力。

VRP系统与其他主要品牌网络交换机的操作系统一样，也提供了用于人——机交互、功能强大的命令行界面（Command Line Interface，CLI）。要使用命令行来配置与管理华为S系列交换机，就必须从认识VRP命令行开始。下面具体介绍。

2.1.2 VRP命令行格式约定

用户可通过在VRP命令行界面下键入文本类配置或管理命令，按下回车键即可把相应的命令提交给网

络交换机执行，从而实现对网络交换机的配置与管理，并可以通过执行相关命令查看输出信息，确认配置结果。与命令行界面（CLI）相对的就是我们通常所说的GUI（Graphical User Interface，图形用户界面），如我们常用的Windows操作系统，是通过鼠标单击相关选项进行设置的。但在CLI下可以一次输入含义更为丰富的指令，系统响应更迅速。

在华为VRP系统中，命令行格式有如表2-1所示的约定。了解这些格式约定，对于理解各个配置或管理命令中的关键字、参数、选项非常重要。本书中的命令行输入格式也遵照表中的约定。

表2-1 VRP命令行格式约定

<table>
<tr><th>格式</th><th>意义</th></tr>
<tr><td>粗体</td><td>命令行关键字（命令中保持不变、必须全部照输的部分）。在命令格式中采用加粗字体表示，但在配置的具体命令中仍为正常体输入，但命令行关键字输入时不区分大小写</td></tr>
<tr><td>斜体</td><td>命令行参数（命令中必须由对应参数的实际值进行替代的部分）。在命令格式中采用斜体表示，但在配置的具体命令中仍为正常体输入</td></tr>
<tr><td>[ ]</td><td>表示为可选参数或可选项，用“[ ]”括起。在命令配置时可以选择，也可以不选择</td></tr>
<tr><td>{ x | y | ... }</td><td>表示二选一，或者多选一参数或选项。在配置命令时必须从中选取其中一个</td></tr>
<tr><td>[x | y | ...]</td><td>表示二选一或多选一可选参数或可选项。在配置命令时可选取其中一个，或者全部不选</td></tr>
<tr><td>{ x | y | ... }*</td><td>表示从二选一（或多）或者多选一(或多)参数或选项。在配置命令时必须从中选取一个或多个，但至少要选取一个，最多可全部选择</td></tr>
<tr><td>[x | y | ...]*</td><td>表示从二选一（或多）或者多选一（或多）可选参数或可选项。在配置命令时可从中选取一个或多个，但也可全部不选</td></tr>
<tr><td>&<1-n></td><td>表示符号&前面的参数可以重复1～n次</td></tr>
<tr><td>#</td><td>表示由“#”开始的行为注释行</td></tr>
</table>

例如authentication-mode { aaa |password }命令是用来设置登录用户界面的验证方式的，其中authentication-mode为命令行关键字；{ aaa |password }中的aaa和password也是命令行关键字，表示从这两个选项中选取一个。再如 vlan batch {vlan-id1 [tovlan-id2] } &<1-10>命令是用来指创建VLAN的，其中的vlan batch为命令行关键字，vlan-id1为参数， to vlan-id2为可选参数，&<1-10>表示前面的VLAN ID或者VLAN ID范围可最多重复10次。

这些不同格式的具体说明在本书后面会全面体现的。

2.1.3 VRP命令行视图

“视图”是VRP命令接口界面，不同的VRP命令需要在不同的视图下才能执行，在不同的视图下也配置有不同功能的命令。但在 VRP 系统中，有的视图又是分层次的，在系统视图下可以进入各种功能视图，在一些功能视图下还可进入对应的子功能视图。

VRP系统的命令行界面分为若干个命令视图，所有命令都注册在某个（或某些）命令视图下。当使用某个命令时，需要先进入这个命令所在的视图。各命令行视图是针对不同的配置要求实现的，它们之间既有联系又有区别。目前在S系列园区交换机中最常见的命令视图、视图功能、提示符示例，以及进入和退出对应视图的方法如表2-2所示（仅列举了部分交换机配置中最常见的视图，不包括全部）。通过quit命令可以返回到上一级命令视图，通过return命令或者按下Ctrl+Z组合键直接返回到用户视图。

表2-2 交换机VRP系统常见命令视图及进入/退出方法

视图	功能	提示符示例	进入命令示例	退出命令
用户视图	查看交换机的简单运行状态和统计信息	<HUAWEI>	与交换机建立连接即进入	**quit**，断开与交换机的连接
系统视图	配置系统参数	[HUAWEI]	在用户视图下键入 **system-view**	**quit** 或 **return**，或 Ctrl+Z 组合键返回用户视图

（续表）

视图	功能	提示符示例	进入命令示例	退出命令
以太网端口视图	配置以太网端口参数	[HUAWEI-Ethernet0/0/1]	百兆以太网端口视图在系统视图下键入 **interface** ethernet 0/0/1	**quit**，返回系统视图；**return**，或 Ctrl+Z 组合键返回用户视图
		[HUAWEI-GigabitEthernet0/0/1]	千兆以太网端口视图在系统视图下键入 **interface** gigabitethernet 0/0/1	
		[HUAWEI-XGigabitEthernet0/0/1]	万兆以太网端口视图在系统视图下键入 **interface** XGigabitEthernet 0/0/1	
NULL 接口视图	配置 NULL 接口视图参数	[HUAWEI-NULL0]	在系统视图下键入 **interface** null 0	
Tunnel 接口视图	配置隧道接口视图参数	[HUAWEI-Tunnel0]	在系统视图下键入 **interface** tunnel 0	
LoopBack 接口视图	配置 LoopBack 接口参数	[HUAWEI-LoopBack0]	在系统视图下键入 **interface** loopback 0	
Eth-Trunk 接口视图	配置 Eth-Trunk 接口参数	[HUAWEI-Eth-Trunk1]	在系统视图下键入 **interface** Eth-Trunk 1	
VLAN 视图	配置 VLAN 参数	[HUAWEI-vlan1]	在系统视图下键入 **vlan** 1	
VLAN 接口视图	配置 VLAN 接口参数	[HUAWEI-Vlanif1]	在系统视图下键入 **interface vlanif** 1	
本地用户视图	配置本地用户参数	[HUAWEI-luser-user1]	在 aaa 视图下键入 **local-user** user1	
VTY 用户界面视图	配置单个或多个 VTY 用户界面参数	[HUAWEI-ui-vty1] 或 [HUAWEI-ui-vty1-3]	在系统视图下键入 **user-interface** vty 1 或 **user-interface** vty 1 3	
Console 用户界面	配置 Console 用户界面参数	[HUAWEI-ui-console0]	在系统视图下键入 **user-interface** console 0	
FTP Client 视图	配置 FTP Client 参数	[ftp]	在用户视图下键入 **ftp** 10.1.1.1	
SFTP Client 视图	配置 SFTP client 参数	sftp-client>	在系统视图下键入 **sftp** 10.1.1.1	
基本 ACL 视图	定义基本 ACL 的子规则（取值范围为 2000～2999）	[HUAWEI-acl-basic-2000]	在系统视图下键入 **acl number** 2000	
高级 ACL 视图	定义高级 ACL 的子规则（取值范围为 3000～3999）	[HUAWEI-acl-adv-3000]	在系统视图下键入 **acl number** 3000	
二层 ACL 视图	定义二层 ACL 的子规则（取值范围为 4000～4999）	[HUAWEI-acl-L2-4000]	在系统视图下键入 **acl number** 4000	**quit**，返回系统视图；**return**，或 Ctrl+Z 组合键返回用户视图
用户自定义 ACL 视图	定义用户自定义 ACL 的子规则（取值范围为 5000～5999）	[HUAWEI-acl-user-5000]	在系统视图下键入 **acl number** 5000	

说明

VRP命令行提示符“HUAWEI”是当前主流5.x版本VRP系统缺省的主机名 （sysname）。通过提示符可以判断当前所处的视图，例如：“<HUAWEI>”表示用户视图，“[HUAWEI] ”表示系统视图，“[HUAWEI-] ”表示系统视图下的其他子视图。有些在系统视图下实现的命令，在其他视图下也可以实现，但实现的功能与命令视图密切相关。

2.1.4 VRP命令级别与用户级别

为了增加交换机的安全性，在VRP系统中把所有命令分成了许多个不同的级别，使不同权限的用户可以使用不同级别的命令。这样也就确定了对应的不同用户级别。不同级别的用户登录后，只能使用等于或低于自己级别的命令。

1．用户级别与命令级别

VRP系统的命令级别分为0～3共4级，但用户级别分成0～15共16个级别。缺省情况下，用户级别和命令级别的对应关系如表2-3所示。

表2-3 VRP用户级别和命令级别对应关系

用户级别	可访问的命令级别	级别名称	可用命令说明
0	0	访问级	网络诊断工具命令（ping、tracert）、从本交换机出发访问外部交换机的命令（也就是作为 Telnet 客户端）、部分 **display** 命令等
1	0、1	监控级	用于系统维护，包括 display 等命令。但并不是所有 **display** 命令都是监控级，比如 **display current-configuration** 命令和 **display saved-configuration** 命令是 3 级管理级
2	0、1、2	配置级	业务配置命令，包括路由、各个网络层次的命令，向用户提供直接网络服务
3～15	0、1、2、3	管理级	用于系统基本运行的命令，对业务提供支撑作用，包括文件系统、FTP、TFTP 下载、用户管理命令、命令级别设置命令、系统内部参数设置命令以及用于业务故障诊断的 **debugging** 命令等。可以通过划分不同的用户级别，为不同管理人员授权使用不同的命令

2．命令级别修改

如果用户需要实现权限的精细管理，可以通过以下两种方法提升某些命令的命令级别。但建议用户不要修改缺省的命令级别，以免造成操作和维护上的不便，甚至给交换机带来安全隐患。

（1）使用 command-privilege level rearrange命令（需要用户确保自己的级别为15级，否则无法执行该命令）将所有缺省注册为2、3级的命令，分别批量提升到10和 15级。命令级别批量提升后，原注册的所有命令行按以下原则自动调整对应的命令级别。

● 0级和 1级命令保持级别不变。

● 2级命令提升到 10级，3级命令提升到 15级。

● 2～9级和11～14级这些命令级别中没有命令。用户可以单独调整需要的命令到这 些级别中，以实现用户权限的精细化管理。

注意

被提升的命令必须是没有被 command-privilege level level view view-name command-key 命令单独修改过的，否则这些命令将维持原来的级别不变。可通过 undo command-privilege level rearrange命令将原来批量提升到10或15级别的命令重新恢复到2 或3 级。但要注意，命令级别恢复的操作对象只能是原来已被command-privilege level rearrange命令批量提升的命令，被 command-privilege level level view view-name command-key 命令单独修改过的命令仍维持原来的级别不变。

但要注意，命令级别批量提升以后，undo command-privilege level rearrange命令本身的级别也被调整到了15级（它原来为3级命令），所以应该确保执行此命令的用户级别是15级。同时，命令级别被批量提升后，会提示用户命令级别已经被批量提升，以减少对其他用户的影响。命令被批量提升后，在没有被恢复之前，如果用户再次执行command-privilege level rearrange命令进行命令级别批量提升操作，此操作将会无效，所有命令的级别都将不会改变。

【示例 1】将所有命令级别批量提升，即将所有缺省注册为2、3级的命令，分别批量提升到10、15级。

<HUAWEI>system-view

[HUAWEI] command-privilege level rearrange

● ou have not set the super password corresponding to a Level 15 user.

It is recommended to quit the operation and set the password.

Are you sure to continue ?[Y / N] y

Info: The Command levels have been upgraded in batch !

（2）使用 command-privilege level level view view-name command-key命令将指定的命令提升到指定的命令级别。命令的参数说明如下。

● level：指定命令新的命令级别，取值范围为0～15的整数；

● view-name：指定要调整命令级别的命令所对应的命令视图名称：shell表示用户视图，system表示系统视图，vlan表示VLAN视图。

● command-key：指定要调整命令级别的具体命令（要是可执行的具体命令，不是带可变值的参数命令），如果命令中包含多个关键字或参数，必须按照关键字或参数的执行顺序依次指定，否则配置无法生效，参数值必须在对应参数取值范围内。

缺省情况下，ping、tracert、telnet等为访问级（0级）；display为监控级（1级）；大部分的配置命令为配置级（2级）；用户密钥设置、FTP、XModem、TFTP以及文件系统操作的命令为管理级（3 级）。使用undo command-privilege [level level] view view-name command-key命令和undo command-privilegeview view-name command-key命令都可以取消当前设置，但是建议用户使用 undo command-privilege view view-name command-key命令格式。

注意

如果修改某视图的某条命令的级别低于缺省级别，请务必注意相应修改quit以 及进入该视图命令的级别。例如system-view、interface命令的缺省级别均为2，如果将命令 interface 开放给级别为 1 的用户使用，则需要通过本命令将 system-view、interface、quit的级别修改为1，以便级别1的用户登录交换机后，能够进入系统视图、接口视图及返回到用户视图。

【示例 2】将reboot命令的用户访问级别提高到15级。

<HUAWEI>system-view

[HUAWEI] command-privilege level 15 view shell reboot

【示例 3】取消设置reboot命令的用户访问级别是15级。

<HUAWEI> system-view

[HUAWEI] user-interface vty0 4

[HUAWEI-ui-vty0-4] user privilege level 15

[HUAWEI-ui-vty0-4] quit

[HUAWEI] undo command-privilege view shell reboot

【示例 4】提升display nqa results命令的级别为 3。

<HUAWEI> system-view

[HUAWEI] command-privilege level 3 view shell display nqa results

【示例 5】降低 interface gigabitethernet 0/0/1命令的级别为 0。

<HUAWEI>system-view

[HUAWEI] command-privilege level 0 view system interface gigabitethernet 0/0/1

3．用户级别的密码设置

用户级别指登录用户的分类，共划分为16个级别（0～15），与命令级别对应，即不同级别的用户登录后，只能使用等于或低于自己级别的命令。

为了防止未授权用户的非法侵入，可以为各个用户级别设置对应的密码，但高级用户访问低级别用户时不需要切换用户级别，也就不需要输入低级别的密码。

可在系统视图下使用 super password [leveluser-level] [cipherpassword] 命令为对应的命令级别设置保护密码。命令中的参数说明如下。

● user-level：可选参数，指定要设置密码的用户级别，取值范围为0～15的整数。缺省情况下是对级别3设置密码。

● password：可选参数，设置对应用户级别的访问密码，为字符串形式，长度范围可以是32位密文密码，如%$%$3W$C.&am)/yJ&EO$0%～>z{ri%$%$；也可以是1～16位的简单口令，如huawei。密码将以加密的形式保存在配置文件中，所以在设置了密码后，无法从系统取回，请妥善保管。如果在命令中不带此可选参数，在直接输入 super password命令或输入 super password leveluser-level命令后会有输入密码的提示，但输入的密码不会在交换机上显示出来。

缺省情况下所有用户级别都没有设置密码，可用undo super password [leveluser-level] 命令取消原来的密码设置。

【示例 6】设置以低级别登录的用户切换到级别10时需要输入密码“123456”。

<HUAWEI>system-view

[HUAWEI] super password level 10 cipher 123456

4．切换用户级别

在从低级别用户切换到高级别用户时，要进行用户身份验证，即需要输入高级别用户密码。方法是在系统视图下使用 super [level] 命令进行操作切换，可选参数 level是用来指定要切换的高用户级别，取值范围为1～15的整数，缺省级别为3，即如果不带此参数，则执行的是切换到用户级别3的操作。

输入该命令后系统将在下面提示输入所要切换到的用户级别的密码，也就是前面介 绍的通过 super password [leveluser-level] [cipherpassword] 命令所设置的对应用户级别的访问密码，并提示你仅可以使用切换后的用户级别，以及比该用户级别更低的所有用户级别的命令。但用户键入的密码不显示在屏幕上，如果3次以内输入正确的密码，则切换到高级别用户，否则保持当前的用户级别不变。

【示例 7】用户切换到最高级别10。

<HUAWEI>super 10

Password:

Now user privilege is 10 level, and only those commands whose level is equal to or less than this level can be used.

Privilege note: 0-VISIT, 1-MONITOR, 2-SYSTEM, 3-MANAGE

2.1.5 VRP命令行编辑

与在PC机命令行输入命令一样，VRP的命令行中也存在一些编辑功能和操作快捷键。另外，在VRP系统命令行中还存在一些操作技巧。了解这些编辑功能、快捷键和操作技巧可大大提高配置与管理命令的输入效率。

1．VRP命令行编辑功能及快捷键

VRP系统的命令行界面提供基本的命令行编辑功能，支持多行编辑，每条命令最大长度为510个字符，且命令关键字不区分大小写，但命令中的参数是否区分大小写则由各自定义的参数而定。用户可以使用系统提供的快捷键，完成对命令的快速输入，从而简化操作。一些常用的编辑功能操作快捷键（都是组合键）如表2-4所示。

表2-4 VRP系统编辑功能键和操作快捷键

功能键或快捷键	说明
普通按键	若编辑缓冲区未满，则插入到当前光标位置，并向右移动光标，否则，响铃告警
退格键 BackSpace	删除光标位置的前一个字符，光标左移
CTRL_A	将光标移动到当前行的开头
左光标键←或 CTRL_B	将光标向左移动一个字符
CTRL_C	停止当前正在执行的功能
CTRL_D	删除当前光标所在位置的字符
CTRL_E	将光标移动到当前行的末尾
右光标键→或 CTRL_F	将光标向右移动一个字符
CTRL_H	删除光标左侧的一个字符
CTRL_K	在连接建立阶段终止呼出的连接
CTRL_N	显示历史命令缓冲区中的后一条命令
CTRL_P	显示历史命令缓冲区中的前一条命令
CTRL_R	重新显示当前行信息
CTRL_T	终止呼出的连接
CTRL_V	粘贴剪贴板的内容
CTRL_W	删除光标左侧的一个字符串（字）
CTRL_X	删除光标左侧所有的字符
CTRL_Y	删除光标所在位置及其右侧所有的字符
CTRL_Z	返回到用户视图
CTRL_]	终止呼入的连接或重定向连接

（续表）

功能键或快捷键	说明
ESC_B	将光标向左移动一个字符串（字）
ESC_D	删除光标右侧的一个字符串（字）
ESC_F	将光标向右移动一个字符串（字）
ESC_N	将光标向下移动一行
ESC_P	将光标向上移动一行

2．VRP命令行编辑操作技巧

有些命令关键字比较长，为了简化输入，VRP系统提供了“不完整关键字输入”功能，即在当前视图下，当输入的字符能够匹配唯一的关键字时，可以不必输入完整的关键字，以提高输入效率和正确性。

比如display current-configuration命令，可以在命令行中输入d cu、di cu或dis cu等都可以执行此命令，但不能输入d c或dis c等，因为以 d c、dis c开头的命令不唯一。

注意

系统可正确执行的命令长度最大为510个字符，包括使用以上介绍的不完整格式 的情况。但如果使用不完整格式进行配置，由于命令保存到配置文件中时使用的是完整格式，可能导致配置文件中存在长度超过510个字符的命令。系统重启时，这类命令将无法恢复。因此，在使用不完整格式的命令进行配置时，也需要注意命令的总长度。

另外，可以在输入不完整的关键字后按下<Tab>键，系统会自动按以下规则补全关键字。

（1）如果与之匹配的关键字唯一，则系统用此完整的关键字替代原输入并换行显示，光标距词尾空一格。

（2）如果与之匹配的关键字不唯一，反复按<Tab>键可循环显示所有以输入字符串开头的关键字，此时光标距词尾没有空格。

（3）如果没有与之匹配的关键字，按<Tab>键后，换行显示，输入的关键字不变。

2.1.6 VRP命令行在线帮助

由于VRP系统命令非常多，许多不常用的命令，或者命令关键字比较长的命令都很难全部记清。这时可以使用VRP系统提供的在线帮助功能，从而无需记忆大量的复杂的命令。

在线帮助通过键入"?"这个特殊的命令来获取，在命令行输入过程中，用户可以随时键入"?"以获得详尽的在线帮助。命令行在线帮助可分为完全帮助和部分帮助。

1．完全帮助

当用户输入命令时，可以使用命令行的完全帮助获取全部关键字或参数的提示。在任一命令视图下，键入"?"获取该命令视图下所有的命令及其简单描述。如在用户视图下输入"？"命令即可显示当前产品的VRP系统的用户视图下所有可用的命令，这时可以通过查看各命令的功能说明或命令本身来确定所需要的命令。具体如下。

```
<HUAWEI> ?
User view commands:
  backup     Backup electronic elabel
  cd         Change current directory
  check      Check information
  clear      Clear information
  clock      Specify the system clock
  compare    Compare function
…
```

还可以在一个命令关键字后面空一个空格后再键入"?"，如果该位置为关键字，则列出全部关键字及其简单描述。下面的示例是在命令后面空一格再加上"？"，提示了两个可接的关键字。

```
<HUAWEI> system-view
[HUAWEI] user-interface vty 0 4
[HUAWEI-ui-vty0-4] authentication-mode ?
  aaa        AAA authentication
  password   Authentication through the password of a user terminal interface
```

如果"？"位置已没有任何参数或关键字，则显示空行。如下所示。

```
[HUAWEI-ui-vty0-4] authentication-mode aaa ?
  <cr>
```

如果在一个命令关键字后面空一个空格后再键入"?"，如果该位置为参数，则列出有关的参数名和参数描述。同样如果"？"位置没有任何关键字或参数，则显示空行。示例如下。

```
<HUAWEI> system-view
[HUAWEI] ftp timeout ?
  INTEGER<1-35791>   The value of FTP timeout, the default value is 30 minutes
[HUAWEI] ftp timeout 35 ?
  <cr>
```

2．部分帮助

当用户输入命令时，如果只记得此命令关键字的开头一个或几个字符，可以使用命令行的部分帮助获取以该字符串开头的所有关键字的提示。可以采用以下几种方式来获取部分帮助。

（1）键入一字符串，其后紧接“?”，即可列出以该字符串开头的所有关键字。示例如下。

```
<HUAWEI>d?
  debugging    delete
  dir          display
```

（2）键入一条命令，在后面接的一字符串后面紧接“?”，即可列出以该字符串开头的所有关键字。示例如下。

```
<HUAWEI>display b?
  bootrom        bpdu
  bpdu-tunnel    bridge
  buffer
```

（3）输入命令的某个关键字的前几个字母，按下<Tab>键，可以显示出完整的关键字，前提是这几个字母可以唯一标示出该关键字，否则，连续按下<Tab>键，可出现不同的关键字，用户可以从中选择所需要的关键字。

2.1.7 VRP命令行的通用错误提示

用户在VRP命令行下面键入任何命令都需要经过语法检查，只有正确才执行，否则系统将会向用户报告错误信息。常见的错误提示信息如表2-5所示。了解这些错误提示 信息所代表的具体含义对于出现一些命令输入或执行错误很有帮助，可以及时发现一些命令输入错误。

表2-5 命令行常见错误提示信息

错误提示信息	错误原因
Error: Unrecognized command found at '^' position.	没有查找到命令，如所输入的命令本身有错误
	没有查找到关键字，如输入的命令关键字不正确
Error: Wrong parameter found at '^' position.	参数类型错，如对应位置本来没有某参数类型，而在你的命令中输入了某参数值
	参数值越界，如你所输入的对应参数的值超出了其取值范围
Error:Incomplete command found at '^' position.	输入命令不完整，如所输入的命令必须要有的关键字或参数没有输入
Error:Too many parameters found at '^' position.	输入参数太多，如所输入的命令中有些参数在命令格式中根本不存在
Error:Ambiguous command found at '^' position.	输入命令不明确

2.1.8 VRP undo命令行

在VRP系统中undo格式命令比较特殊，几乎所的配置命令（不包括管理类的命令）都有对应的undo命令格式，其中undo作为这些命令的关键字，即为undo命令行。

undo 命令行一般用来恢复缺省情况、禁用某个功能或者删除某项设置。但大多数undo 命令行的作用就是用来恢复对应命令的缺省设置。如在前面介绍的 super password [leveluser-level] [cipher password] 命令是用来设置对应用户级别的访问密码的，如果要恢复对应用户级别的缺省无密码设置，即删除原来所设置的密码，则可使用undo super password [level user-level] 命令。但在这里要注意的是，undo命令行格式有多

种，下面分别介绍。

1．不带有原命令中的参数和选项

有的undo命令行只需在原命令的关键字前面加上undo关键字，后面的参数和选项都不用带，如sysname host-name的undo命令行undo sysname，authentication-mode { aaa |password }的 undo命令行undo authentication-mode。这类命令通常是一些最终仅可带一个参数或选项（即二选一，或多选一参数或选项），上面两条命令都属于这种类型；或者根本不带参数和选项的命令，主要是一些功能启用命令，如启用 telnet 服务的telnet server enable命令对应的 undo命令行即为undo telnet server enable。

2．仅带原命令中前面的部分参数或选项

也有一些命令的 undo 命令行是需要带有部分参数和选项的，如前面说的 super password [leveluser-level] [cipherpassword] 命令的 undo命令行undo super password [leveluser-level]，它只带了用户级别这个参数，后面的密码设置参数没有带。这类命令通常是带有多个包括关键字的参数、选项的命令，但这些参数、选项不是并列的，通常前面的参数或选项是主体，后面的参数或选项设置为作用在前面的主体参数或选项之上的，这时在取消设置时仅需要指出最前面一个或者多个参数或选项。如前面介绍的命令中 user-level 参数是后面 password 参数的主体，所以在其 undo 命令行没有包括 cipher password可选参数。

3．带有原命令中全部的参数和选项

还有一些命令的undo命令行是需要带有全部的参数和选项，这些命令通常是带有多个并列的参数或选项，要删除设置时必须全部指定各参数的取值。如用来批量创建VLAN的 vlan batch { vlan-id1 [to vlan-id2] } &<1-10>命令所对应的 undo命令行为undo vlan batch { vlan-id1 [to vlan-id2] } &<1-10>，就带有了原命令中的全部参数了。这类命令的undo命令行如果不全部是原命令的参数、选项，则在命令执行时系统无法确认所要恢复或删除的参数值。如这里的undo vlan batch命令不带任何参数的话，理论上来讲就是要删除所有VLAN，但事实上在大多数情况下不能这样操作，很危险；如果仅带了vlan-id1参数，则仅会删除对应的一个VLAN，如果想要删除一个范围的VLAN，必须同时带上to vlan-id2参数，如果要删除多个不连续范围的VLAN，则还要同时指出由&<1-10>决定的其他VLAN ID范围。

2.1.9 查看历史命令

VRP命令行界面能够自动保存用户键入的历史命令。当用户需要输入之前已经执行过的命令时，可以调用命令行界面保存的历史命令，并重复执行，方便用户的操作。

缺省情况下，为每个登录用户保存 10 条历史命令。可以通过 history-command max-size size-value命令在相应的用户界面视图下重新设置保存历史命令的条数，最大设置为 256。但不推荐用户将此值设置过大，因为可能会花费较长时间才查看到所需要的历史命令，反而影响了效率。

对历史命令的操作方法如表2-6所示。

表2-6 历史命令访问操作方法

操作任务	命令或功能键	结果
显示历史命令	**display history-command [all-users]**	不指定 **all-users** 可选项时，则显示当前用户键入的历史命令；指定 **all-users** 可选项时，则显示得是所有登录用户键入的历史命令
访问上一条历史命令	上光标键或者<Ctrl_P>组合键	如果还有更早的历史命令，则取出上一条历史命令，否则响铃警告
访问下一条历史命令	下光标键或者<Ctrl_N>组合键	如果还有更新的历史命令，则取出下一条历史命令，否则显示为空，响铃警告

说明

对于Windows 9x系统（现在已很少有人用了）的超级终端，↑光标键无效，这是 由于Windows 9x的超

级终端对这个键作了不同解释，这时可以用<Ctrl_P>组合键代替↑光标键达到同样目的。

另外，保存的历史命令与用户输入的命令格式相同，如果用户使用了命令的不完整形式，保存的历史命令也是不完整形式。如果用户多次执行同一条命令，则历史命令中只保留最近的一次。但如果执行时输入的形式不同，将作为不同的命令对待。例如：多次执行 display ip routing-table命令，历史命令中只保存最近这一条。如果分别执行display ip routing-table（完整格式）和 dis ip rout（不完整格式），将保存为两条历史命令。

2.2 查看命令行显示信息

我们在交换机的VRP命令行中配置了许多命令，如果想要查看以往的配置命令信息又将如何进行呢？本节将具体介绍查看命令行显示信息，包括查询命令行的配置信息、控制命令行显示方式和过滤命令行显示信息3个方面。

2.2.1 查询命令行的配置信息

在完成一系列配置后，可以执行相应的 display 命令查看交换机的配置信息和运行信息。比如，在完成了FTP服务功能的各项配置后，可以执行display ftp-server命令查看当前FTP服务器的各项参数；完成了VLAN方面的各项配置后，可以执行display vlan命令查看所有的VLAN相关配置信息；完成了STP方面的各项配置后，可以执行display stp命令查看STP相关配置信息。

注意

在这些 display 命令的输出信息中，对于某些正在生效的配置参数，如果某参数 的设置值采用了缺省值则不会在输出信息中显示。如果要同时显示当前视图下未被修改的缺省配置，可以执行命令 display this include-default进行查看。另外，对于某些参数，虽然用户已经配置，但如果这些参数所在的功能并没有生效，则也不会在输出信息中显示。

1．查看当前生效的配置信息

VRP系统还支持对当前生效的所有配置信息和当前视图下的所有配置信息分别查看。执行display current-configuration [configuration [configuration-type [configuration-instance]] |interface [interface-type [interface-number]]] [feature feature-name [filter filter-expression] | filter filter-expression] 或display current-configuration [all | inactive] 命令即可查看当前生效的所有配置信息，也可通过其中的参数或关键字查看指定的配置类型、配置实例、接口或特性等的配置信息，或由过滤条件，或者由正则表达式过滤要显示的配置信息。各部分是以“#”行分隔的。有关正则表达式将在2.2.3节具体介绍。以上两命令中的参数和选项说明如下。

（1）configuration-type：多选一可选参数，显示指定的配置类型的配置（但依赖于系统当前已有的配置），如可显示AAA配置、系统配置、用户界面配置等。

（2）configuration-instance：可选参数，显示指定的 VPN 配置实例中的配置，VPN实例名为1～80个字符。

（3）interface-type [interface-number]：多选一可选参数，显示指定接口的配置。

（4）feature-name：多选一可选参数，显示指定特性的配置。

（5）filter-expression：可选参数，指定用于过滤配置信息的过滤表达式，为 1～255个字符，不支持空格，不区分大小写。

（6）all：二选一选项，指定显示所有板卡的配置信息，包括不在位的板卡的配置信息。

（7）inactive：二选一选项，指定显示不在位的板卡的配置信息。

【示例 1】查看包含字符串“vlan”的所有配置信息。这是通过正则表达式来过滤显示信息的，有关正则表达式将在2.2.3节介绍。

```
<HUAWEI>display current-configuration | includevlan
vlan batch 10 77 88
port link-type trunk
port trunk allow-pass vlan 10
```

【示例 2】查看ftp特性的配置信息。

```
<HUAWEI>display current-configuration feature ftp
#
FTP server enable
#
------------ END ------------
```

2．查看当前视图下正在运行的配置信息

可通过display this命令查看当前视图下正在运行的配置信息。当用户在某一视图下完成一组配置之后，需要验证是否配置正确，则可以执行本命令，但这仅显示当前视图下的生效配置。

2.2.2 控制命令行显示方式

VRP所有的命令行有共同的显示特征，并且可以根据用户的需求，灵活控制显示方式。命令行的回显模式（也就是在屏幕上的显示模式）分为字符模式和行模式，可通过terminal echo-mode { character | line }命令设置，缺省情况下为字符模式。如果设置为character 模式，则指定命令行的回显模式是字符模式。此时输入命令行时，用户输入一个字符系统显示一个字符；如果设置为line模式，则指定命令行的回显模式是行模式。此时输入命令行时，用户输入字符后，只有在按下回车键，或者<Tab>键或？键，系统才回显输入的字符，这可提高命令输入时的安全性，因为前面输入的字符都看不到。

当终端屏幕上显示的信息过多时，可以使用<PageUp>和<PageDown>键显示上一页信息和下一页信息。当执行某一命令后，如果显示的信息超过一屏时，系统会自动暂停，以方便用户查看。此时用户可以通过功能键控制命令行的显示方式，如表2-7所示。当然，也可以事先通过 screen-length screen-length temporary 命令设置当前终端屏幕的临时显示行数，如果参数 screen-length 的取值为 0，则关闭分屏功能，即当显示的信息超过一屏时，系统不会自动暂停。

表2-7 命令行显示方式的控制方式

功能键	功能
按下<Ctrl_C>或<Ctrl_Z>组合键	停止显示或命令执行。也可以键入除空格键、回车键等的其他键（可以是数字键或字母键）停止显示和命令执行
键入空格键	继续显示下一屏信息
键入回车键	续续显示下一行信息

2.2.3 过滤命令行显示信息

在通过VRP系统中的display命令查看显示信息时，可以使用正则表达式（即指定显示规则）来过滤显示信息。过滤命令行显示信息可以帮助用户迅速查找到所需要的信息。

过滤命令行显示信息的使用方法有以下两种。

（1）在命令中指定过滤方式：在命令行中通过输入begin、exclude或include关键字加正则表达式的方式来过滤显示。begin 关键字是显示特定行和其以后的所有行，该特定行必须包含指定正则表达式；exclude关键字用来显示不包含指定正则表达式的所有行；include关键字用来指定只显示包含指定正则表达式的所有行。

（2）在分屏显示时指定过滤方式：在分屏显示时，使用“/”、“－”或“+”符号加正则表达式的方式，可以对还未显示的信息进行过滤显示。其中，“/”等同关键字begin；“－”等同关键字exclude；“+”等同关键字include。

1．通过正则表达式过滤

正则表达式描述了一种字符串匹配的模式，由普通字符（例如字符 a 到 z）和特殊字符（或称“元字符”）组成。正则表达式作为一个模板，将某个字符模式与所搜索的字符串进行匹配。

正则表达式一般具有以下功能。

（1）检查字符串中符合某个规则的子字符串，并可以获取该子字符串。

（2）根据匹配规则对字符串进行替换操作。

（3）正则表达式由普通字符和特殊字符组成。

普通字符匹配的对象是普通字符本身，包括所有的大写和小写字母、数字、标点符号以及一些特殊符号。例如：a匹配abc中的a，202匹配202.113.25.155中的202，@匹配xxx@xxx.com中的@。正则表达式为1～255个字符的字符串，区分大小写。它还支持多种特殊字符，特殊字符的匹配规则如表2-8所示。

表2-8 特殊字符及其语法意义描述

<table>
<tr><th>特殊字符</th><th>含义</th><th>使用说明</th></tr>
<tr><td>^string</td><td>行首匹配符，string 只能出现在每行的开始</td><td>如“^user”只能匹配以 user 开始的行，不能匹配以像 Auser 或者其他字符开始的行</td></tr>
<tr><td><i>string</i>$</td><td>行尾匹配符，<i>string</i> 只能出现在每行的末尾</td><td>如“user$”只能匹配以 user 结尾的行，不能匹配以像 userA 或者其他字符结尾的行</td></tr>
<tr><td>.</td><td>句点，通配符，匹配任何一个字符，包括单个字符、特殊字符和空格等</td><td>如“.l”（这是 L 的小写，不是数字 1）可以配置 vlan 和 mpls 等</td></tr>
<tr><td>*</td><td>星号，匹配“*”前面的字符或字符组零次或多次</td><td>如“zo*”可以匹配 z（匹配前面的字符）和 zoo（匹配前面的字符组）；(zo)*（此时仅可以匹配字符组）可以匹配 zo 和 zozo</td></tr>
<tr><td>+</td><td>加号，匹配“+”前面的字符或字符组一次或多次</td><td>如“zo+”可以匹配 zo 和 zoo，但不能匹配 z</td></tr>
<tr><td>|</td><td>竖线，匹配“|”左边的整个字符串或者右边的整个字符串</td><td>如“def|int”只能匹配包含 def 或 int 的字符串</td></tr>
<tr><td>_</td><td>下划线，该字符出现在表达式的开头或结尾时，等效于行首匹配符或行尾匹配符（即特殊字符^或$），其他情况下等效于逗号、空格或者作为普通字符时的左括号、右括号、左大括号、右大括号</td><td>如“a_b”可以匹配“a b”和“a(b”等；而“_ab”却只能匹配以 ab 开头的行；“ab_”只能匹配以 ab 结束的行</td></tr>
<tr><td>-</td><td>连接符，用于连接两个数值或字母（小的在前，大的在后），与“[ ]”符号连用表示一个范围</td><td>如从 1 到 9 表示为 1-9（包括 1 和 9）；从 a 到 h 表示为 a-h（包括 a 和 h）</td></tr>
</table>

（续表）

特殊字符	含义	使用说明
[ ]	表示字符选择范围，将以选择范围内的单个字符为条件进行匹配，只要字符串里包含该范围的某个字符就能匹配到	如“[16A]”表示可以匹配到的字符串只需要包含 1、6 或 A 中任意一个；“[1-36A]”表示可以匹配到的字符串只需要包含 1、2、3、6 或 A 中任意一个（-为连接符） 如果“]”需要作为普通字符出现在[ ]内时，必须把“]”写在[ ]内字符的最前面，形如“[]string]”才能匹配到，而“[”没有这样的限制
( )	表示字符组，一般与“+”或“*”等符号一起使用	如“(123A)”表示字符组 123A；“408(12)+”可以匹配 40812 或 408121212 等字符串，但不能匹配 408
\index	表示重复一次指定字符组，字符组是指\前用()括起来的字符串，index 对应\前字符组的顺序号按从左至右的顺序从 1 开始编号：如果\前面只有一个字符组，则 index 只能为 1；如果\前面有 *n* 个字符组，则 index 可以为 1～*n* 中的任意整数	如“(string)\1”表示把 string 重复一次，匹配的字符串必须包含 stringstring；“(string1)(string2)\2”表示把 string2 重复一次，匹配的字符串必须包含 string1string2string2；“(string1)(string2)\1\2”表示先把 string1 重复一次，再重复一次 string2，匹配的字符串必须包含 string1string2string1string2
[^]	表示选择范围外的字符，将以单个字符为条件进行匹配，只要字符串里包含该范围外的某个字符就能匹配到	如“[^16A]”表示可匹配的字符串只需要包含 1、6 和 A 之外的任意字符，该字符串也可以包含字符 1、6 或 A，但不能只包含这 3 个字符。比如“[^16A]”可以匹配 abc、m16，不能匹配 1、16、16A
\<*string*	匹配以 *string* 开头的字符串	如“\<do”可以匹配单词 domain，还可以匹配字符串 doa
string\>	匹配以 *string* 结尾的字符串	如“do\>”可以匹配单词 undo，还可以匹配字符串 abcdo
\b*character2*	匹配 *character1character2*，*characte1* 可以是除了数字、字母和下划线外的任意字符，\b 等效于[^A-Za-z0-9_]	如“\ba”可以匹配−a，−为 *character1*，a 为 *character2*，但是不能匹配 2a 和 ba 等
\B*character*	匹配到的字符串中必须包含字符 *character*，且 *character* 前不能是空格	如“\Bt”可以匹配 install 中的 t，而不能匹配 big top 中的 t
character1\w	匹配 *character1character2*，*character2* 必须是数字、字母或下划线，\w 相当于[A-Za-z0-9_]	如“v\w”能匹配到 vlan，v 为 *character1*，l 为 *character2*，v\w 还能匹配 service，i 为 *character2*
\W	等效于\b	如“\Wa”可以匹配−a，−为 *character1*，a 为 *character2*，但是不能匹配 2a 和 ba 等
\	转义操作符，\后紧跟本表列的单个特殊字符时，将去除特殊字符的特定含义	如“\\”可以匹配包含\的字符串，“\^”可以匹配包含^的字符串，“\\b”可以匹配包含\b 的字符串

在实际应用中，往往不是一个普通字符加上一个特殊字符配合使用，而是由多个普通字符和特殊字符组合，匹配某些特征的字符串。

说明

某些特殊字符如果处在如下的正则表达式的特殊位置时，会引起退化，这些特殊 字符又将成为普通字符。

● 特殊字符处在转义符号‘\’之后，则发生转义，变为匹配该字符本身。

● 特殊字符“*”、“+”、“?”，处于正则表达式的第一个字符位置。例如：+45 匹配+45，abc(*def)匹配abc*def。

● 特殊字符“^”，不在正则表达式的第一个字符位置。例如：abc^匹配abc^。

● 特殊字符“$”，不在正则表达式的最后一个字符位置。例如：12$2匹配12$2。

● 右括号“)”或者“] ”没有对应的左括号“(”或“[”。例：abc] 匹配 abc)，0-9)匹配0-9)。

2．在命令中指定过滤方式

华为交换机可采用正则表达式实现管道符“|”的过滤功能，但并非所有 display 命令均支持管道符。当显示信息内容很多时，此 display 命令支持管道符；当显示信息内容很少时，此display命令不支持管道符。

按过滤条件进行查询时，显示内容的第一行信息中以包含该字符串的整条信息作为起始。在支持正则表达式的命令中，有3种过滤方式可供选择。

● |begin regular-expression：输出以匹配指定正则表达式的行开始的所有行。即过滤掉所有待输出字符串，直到出现指定的字符串（此字符串区分大小写）为止，其后的所有字符串都会显示到界面上。

● | exclude regular-expression：输出不匹配指定正则表达式的所有行。即待输出的字符串中如果没有包含指定的字符串（此字符串区分大小写），则会显示到界面上；否则过滤不显示。

● | include regular-expression：只输出匹配指定正则表达式的所有行。即待输出的字符串中如果包含指定的字符串（此字符串区分大小写），则会显示到界面上；否则过滤不显示。

【示例 1】执行命令display interface brief，显示不匹配正则表达式“Ethernet|NULL|Tunnel”的所有行。其中的管道符中包括的“Ethernet|NULL|Tunnel”表示只要匹配“Ethernet”、“NULL”或“Tunnel”其中一个即不显示。

```
<HUAWEI>display interface brief | exclude Ethernet|NULL|Tunnel
PHY: Physical
*down: administratively down
^down: standby
(l): loopback
(s): spoofing
(b): BFD down
(e): ETHOAM down
(dl): DLDP down
(d): Dampening Suppressed
InUti/OutUti: input utility/output utility
Interface      PHY   Protocol InUti  OutUti   inErrors   outErrors
Eth-Trunk1   down   down   0%   0%   0   0
Eth-Trunk17   down   down   0%   0%   0   0
LoopBack1   up   up(s)   0%   0%   0   0
Vlanif1   up   down   --   --   0   0
MEth0/0/1   down   down   0%   0%   0   0
Vlanif2   down   down   --   --   0   0
Vlanif10   down   down   --   --   0   0
Vlanif12   down   down   --   --   0   0
Vlanif13   down   down   --   --   0   0
Vlanif20   up   up   --   --   0   0
Vlanif22   down   down   --   --   0   0
Vlanif222   down   down   --   --   0   0
Vlanif4094   down   down   --   --   0   0
```

【示例 2】执行命令display current-configuration，只显示匹配正则表达式“vlan”的所有行。

```
<HUAWEI>display current-configuration | include vlan
vlan batch 2 10 101 to 102 800 1000
vlan 2
vlan 10
port trunk pvid vlan 800
```

undo port trunk allow-pass vlan 1
port trunk allow-pass vlan 10 101 800
undo port hybrid vlan 1
undo port hybrid vlan 1
port hybrid untagged vlan 10
undo port hybrid vlan 1
undo port hybrid vlan 1

3．在分屏显示时指定过滤方式

支持在分屏显示时指定过滤方式的命令行有：

● display current-configuration
● display interface
● display arp

采用分屏显示时，可以在分屏提示符“---- More ----”中指定以下过滤类型：

● /regular-expression：输出以匹配指定正则表达式的行开始的所有行。
● －regular-expression：输出不匹配指定正则表达式的所有行。
● +regular-expression：只输出匹配指定正则表达式的所有行。

2.3 VRP文件系统管理

文件系统管理就是用户对交换机中存储的文件和目录的访问管理，如用户可以通过命令行对文件或目录进行创建、移动、复制、删除等操作，并可对交换机存储器进行管理。它们都是在用户视图下进行的。VRP系统是基于Linux操作系统平台进行二次开发的，所以它的文件系统管理命令和操作方法与我们常用的Linux系统中的对应操作方法完全一样（其实许多命令也与早期的DOS系统是一样的）。

2.3.1 VRP文件系统概述

华为S系列交换机上的所有文件（如配置文件、系统软件等）都是以VRP文件系统的方式被有效地管理。VRP 文件系统实现两类功能：管理存储器（包括 flash:和 cfcard:存储器）和管理保存在存储器中的文件。

1．VRP系统文件名格式

VRP系统中的文件名都是字符串形式，长度范围是1～160，不区分大小写。文件名 有两种表示方式：文件名、路径+文件名。如果直接使用这种文件名，则表示当前工作路径下的文件。

如果文件不在当时工作路径下，则格式为 drive + path + filename，直接指定到某路径下的文件名。其中drive是交换机中的存储器，命名为flash：或cfcard：。如果交换机在堆叠情况下，则drive的命名如下。

● flash：堆叠系统中主交换机Flash存储器根目录。
● 槽位号#flash：堆叠系统中某槽位号的Flash存储器根目录。

例如：slot2#flash:是指槽位号2的Flash卡。

path是指存储器中目录以及子目录，即路径。目录名使用的字符不可以是空格、“～”、“*”、“/”、“\”、“:”、“'”、“"”等字符，不区分大小写。

2．绝对路径与相对路径

交换机支持的路径可以是绝对路径也可以是相对路径。指定根目录（指定drive）的路径是绝对路径，

相对路径有相对于根目录（即当前的存储器目录）的路径和相对于当前工作路径的路径，路径以“/”开头，则表示相对于根目录的路径。

若路径为“flash:/my/test/”，这是绝对路径；若路径为“/selftest/”，表示为根目录下的selftest目录，这就是相对于根目录的相对路径；若路径为“selftest/”，表示当前路径下的selftest目录，这是相对于当前路径的相对路径。

例如用dir flash:/my/test/mytest.txt命令查看flash:/my/test/路径下的mytest.txt文件的信息，这是一种绝对路径表示方法。如果要用相对于根目录的路径来表示，则可以使用以下命令：dir/my/test/mytest.txt；如果要用相对于当前路径的路径（假设当前工作路径为 flash:/my/），则使用dir test/mytest.txt命令。

2.3.2 目录管理

当需要在客户端与服务器端进行文件传输时，需要使用文件系统对目录进行配置。可以使用表2-9中的用户视图命令来进行相应的目录操作，包括创建或删除目录、显示当前的工作目录、指定目录下文件或目录的信息等。

表2-9 VRP系统目录操作命令

目录操作	所用命令	说明
创建目录	**mkdir** *directory*	创建指定目录，但所创建的目录名不能与指定目录下的其他目录或文件名重名。参数 *directory* 用来指定要创建的目录（包括路径），长度为 1～64 个字符。建议采用“驱动器名” ＋ “:” ＋ “/” ＋ “目录名” 的组合。其中目录名使用的字符不可以是空格、“～”、“*”、“/”、“\”、“:”、“'”、“"” 等字符，不区分大小写。**如果不指定目录路径，则代表当在前目录下创建**
删除目录	**rmdir** *directory*	删除指定目录。参数 *directory* 用来指定要删除的目录（包括路径），其他说明同上面介绍的 **mkdir** 命令的该参数说明。**如果不指定目录路径，则代表当在前路径下删除指定的目录** 所删除的目录必须为空目录，否则将无法进行操作；另外，执行本命令后，在回收站中的原来属于该目录中的文件也会被自动删除

（续表）

<table>
<tr><th>目录操作</th><th>所用命令</th><th>说明</th></tr>
<tr><td>显示
当前路径</td><td>pwd</td><td>仅用来显示当前所处的目录路径信息</td></tr>
<tr><td>进入指定
的目录</td><td>cd directory</td><td>修改当前工作路径或切换至其他存储器交换机的目录。参数 directory 用来指定要进入的目标目录名，其他说明同上面介绍的 mkdir 命令的该参数说明，例如：cfcard:/selftest/test/</td></tr>
<tr><td>显示
目录或
文件信息</td><td>S2700/3700 系列交换机：
dir [/all] [filename | flash:]
S5700/6700/7700/9700 系列交换机：
dir [/all] [filename | directory | /all-filesystems]</td><td>查看存储器中指定的文件和目录的信息，支持通配符“*”。命令中的参数和选项说明如下。
（1）/all：可选项，指定查看当前路径下的所有的文件和目录，包括已经放入回收站的文件。在回收站中的文件名用“[ ]”标识
（2）filename：可选参数，指定要显示的文件名称，为 1～160 个字符。建议采用“驱动器名” ＋ “:” ＋ “/” ＋ “目录名” ＋ “/” ＋ “文件名” 的组合。其中目录名使用的字符不可以是空格、“～”、“*”、“/”、“\”、“:”、“'”、“"” 等字符，不区分大小写
（3）directory：多选一可选参数，指定要显示的目录路径，其他说明同上面介绍的 mkdir 命令的该参数说明
（4）flash:：二选一可选项，表示查看闪存根目录下的所有文件和目录
（5）/all-filesystems：多选一可选项，指定显示交换机上所有存储器根目录中文件和目录的信息</td></tr>
</table>

说明

表中的“驱动器名”，在S2700、S3700、S5700和S6700系列中仅指闪存“flash：”（这个冒号不能少），但在 S770 和 S9700 系列中，除了闪存外（但分主、为主控板闪存“flash：”和备用主控板闪存“slave#flash：”），还包括主控板 CF 卡“cfcard：”和备用主控板CF卡“slave#cfcard：”；如果堆叠交换机，则驱动器名应为“框号/槽位号#cfcard：”或“框号/槽位号#flash：”

这里的路径可以是绝对路径也可以是相对路径。相对路径有相对于根目录（即当前的存储器目录）的路径和相对于当前工作路径的路径，路径以“/”开头，则表示相对于根目录的路径。如路径为“cfcard:/my/test/”，表示绝对路径；如路径为“/selftest/”，表示根目录下的selftest目录，这是相对于根目录的相对路径；如路径为“selftest/”，表示当前工作路径下的selftest目录，这是相对于当前工作路径的相对路径。

以上说明同样适用于下节将要介绍的VRP文件系统管理。

【示例 1】在当前flash:存储器目录下创建子目录cfg。执行命令后会有提示信息提示该目录创建成功。

```
<HUAWEI>mkdir flash:/cfg
Info: Create directory flash:/cfg. . .Done.
```

【示例 2】删除当前路径（主控板CF卡根目录）下的test子目录。执行命令后会有提示信息要求你再一次确认，确认后进行删除。删除成功后也会有提示信息。

```
<HUAWEI>rmdir test
Remove directory cfcard:/test?[Y/N]:y
%Removing directory cfcard:/test. .Done!
```

【示例 3】显示当前工作路径。从输出可以看出，当前的工作路径是在闪存的根目录下。

```
<HUAWEI>pwd
flash:
```

【示例 4】从当前的闪存根目录进入闪存下的test 目录中。可以先通过pwd命令查看当前路径，进入后同样可以使用pwd命令验证当前路径是否修改成功。

```
<HUAWEI>pwd
flash:
<HUAWEI> cd test
<HUAWEI>pwd
flash:/test
```

【示例 5】查看当前路径下的test.bak文件信息。

```
<HUAWEI> dir test.bak
Directory of flash:/
0    -rw-    11779    Apr 05 2006 10:23:03    test.bak
31877 KB total (15961 KB free)
```

【示例 6】查看当前路径下的所有文件的目录信息。

```
<HUAWEI>dir /all
Directory of flash:/
  Idx    Attr    Size(Byte)    Date    Time    FileName
  0    -rw-    889    Feb 25 2012 10:00:58    private-data.txt
  1    -rw-    6,311    Feb 17 2012 14:05:04    backup.cfg
  2    -rw-    836    Jan 01 2012 18:06:20    rr.dat
```

3 drw- - Jan 01 2012 18:08:20 syslogfile
4 -rw- 836 Jan 01 2012 18:06:20 rr.bak
5 drw- - Feb 27 2012 00:00:54 resetinfo
6 -rw- 523,240 Mar 16 2011 11:21:36 bootrom_53hib66.bin
7 -rw- 2,290 Feb 25 2012 16:46:06 vrpcfg.zip
8 -rw- 812 Dec 12 2011 15:43:10 hostkey
9 drw- - Jan 01 2012 18:05:48 compatible
10 -rw- 25,841,428 Nov 17 2011 09:48:10 s-sbox_l3b070.cc
11 -rw- 540 Dec 12 2011 15:43:12 serverkey
12 -rw- 26,101,692 Dec 21 2011 11:44:52 s-sbox_l3b120.cc
13 -rw- 6,292 Feb 14 2012 11:14:32 1.cfg
14 -rw- 6,311 Feb 17 2012 10:22:56 1234.cfg
15 -rw- 6,311 Feb 25 2012 17:22:30 [11.cfg]
65,233 KB total (13,632 KB free)

说明

在以上输出信息中，第一列“Idx”代表文件或目录的索引，或者称序号，第二列 “Attr”指文件或目录属性。它分为四部分，第一部分表示是文件还是目录，如果是目录则显示“d”，如果是文件则显示“-”；后面三部分均表示当前用户对该目录或文件所具有的访问权限，r表示可读，w表示可写，x表示可执行。因为是在最低级别的用户视图下执行，所以均没有 x（可执行）权限。如果文件名或目录名用“[ ] ”括住了，则表示该文件或目录是在当前存储器的回收站中的，如上面的[11.cfg] 。

2.3.3 文件管理

可以使用表 2-10中的用户视图命令（但其中的 execute和 file prompt命令需要在系 统视图下执行）对华为S系列交换机软件系统进行相应的文件操作，包括删除文件、重命名文件、复制文件、移动文件、查看文件的内容、显示指定文件的信息等。

表2-10 文件管理命令

文件操作	所用命令	说明
显示文本文件内容	**S2700/3700 系列交换机**： **more** *filename* **S5700/6700/7700/9700 系列交换机**： **more** *filename* [*offset*] [**all**]	显示指定文件内容。命令中的参数和选项说明如下。 （1）*filename*：指定待显示文件的路径和文件名。其他说明同表 2-9 中介绍的 **dir** 命令的该参数说明 （2）*offset*：可选参数，指定待显示文件的偏移量，取值范围是（0～2147483647）整数个字节 （3）**all**：可选项，指定一次显示文件内的全部内容，不进行分屏显示
复制文件	**copy** *source-filename destination-filename* [**all**]	把源文件复制为目标文件，支持通配符"*"。命令中的参数和选项说明如下。 （1）*source-filename*：指定被复制文件的路径名或源文件名，其他说明同表 2-9 中介绍的 **dir** 命令的 *filename* 参数说明 （2）*destination-filename*：目标文件的路径或路径及目标文件名，其他说明同上面介绍的 **dir** 命令的 *filename* 参数说明 **【说明】**如果目标文件的目录路径与源文件的目录一致，则目标文件的目录路径可省略；如果目标文件名与源文件一样，则目标文件名可省略；如果目标文件名与已经存在的文件重名，会提示是否覆盖，操作成功后原有同名文件将被覆盖；如果只指定目标文件的路径，而没有指定目的文件名称，则缺省是使用源文件名作为目标文件名，但是如果目标文件和被复制文件在一个目录下，必须指定目标文件的文件名，否则复制将不成功 （3）**all**：可选项，复制文件到所有堆叠成员交换机上。此可选项仅可在堆叠交换机上使用
移动文件	**move** *source-filename destination-filename*	将源文件从指定目录移动到目标目录中，移动时有确认提示。参数 *source-filename* 用来指定被移动的源文件的路径和文件名，参数 *destination-filename* 用来指定目标文件的路径和文件名，其他说明同表 2-9 中介绍的 **dir** 命令的 *filename* 参数说明。**但此命令执行的源文件和目标文件必须在相同的存储器下，否则系统会报错** 如果目标文件名与已经存在的文件重名，操作成功后原有同名文件将被覆盖；如果只指定目标文件的路径，而没有指定目标文件名称，则缺省是使用源文件名作为目标文件名
重命名目录或文件	**rename** *old-name new-name*	对目录或文件进行重命名，重命名时有确认提示。参数 *old-name* 用来指定当前目录名或文件名；*new-name* 用来指定重命名后的目录名或文件名，其他说明参见表 2-9 中 **mkdir** 命令中的参数 *directory* 说明 该命令不支持跨路径的文件重命名，即重命名的源目录和目标目录、源文件和目标文件必须在同一路径下；且如果目标文件名与已经存在的目录名重名，或者目标文件名与已经存在的文件名重名，都将出现错误提示信息

（续表）

<table>
<tr><th>文件操作</th><th>所用命令</th><th>说明</th></tr>
<tr><td>压缩文件</td><td>zip source-filename
destination- filename</td><td>压缩指定文件（压缩后的文件名可以不一样）。但要注意，这里压缩后文件大小不仅不会变小，还可能变大，只是生成了压缩格式文件，便于备份。参数 source-filename 用来指定被压缩的源文件名；参数 destination-filename 用来指定压缩后的目标文件名，其他说明同表 2-9 中介绍的 dir 命令的 filename 参数说明。压缩后的文件扩展名为.zip
如果只指定了目标文件所在的路径，但未指定目标文件名，则目标文件名与源文件名相同。压缩后，源文件仍然存在。但只能对文件进行压缩，不能压缩目录</td></tr>
<tr><td>解压缩
文件</td><td>unzip source-filename
destination- filename</td><td>解压缩指定文件（解压缩后的文件名可以不一样）。参数 source-filename 用来指定被解压缩的源文件名；参数 destination-filename 用来指定解压缩后的目标文件名，其他说明同表 2-9 中介绍的 dir 命令的 filename 参数说明。如果只指定了目标文件所在的路径，但未指定目标文件名，则目标文件名与源文件名相同。解压缩后，源文件仍然存在
压缩文件的类型必须是.zip 的压缩文件，如果是其他类型文件，在解压缩过程中系统会提示出错。且压缩文件的源文件必须是单个文件，如果是一个目录或者多个文件可能会导致解压缩失败</td></tr>
<tr><td>删除文件</td><td>S2700/3700 系列交换机：
delete [/unreserved]
filename [all]
S5700/6700/7700/9700 交换机：
delete [/unreserved]
[/quiet] { filename | devicename }</td><td>删除指定文件，支持通配符“*”。命令中的参数和选项说明如下。
（1）/unreserved：可选项，指定被删除的文件不可以使用 undelete 命令恢复
（2）filename：指定要删除的文件的路径和文件名。其他说明同表 2-9 中介绍的 dir 命令的 filename 参数说明
（3）/quiet：可选项，指定无需确认直接删除文件。此可选项要慎用，因为在删除过程中不会再有确认提示了
（4）all：可选项，指定批量删除所有成员交换机上对应路径下的文件，仅在堆叠交换机上可用</td></tr>
<tr><td>恢复
回收站
中的文件</td><td>S2700/3700 系列交换机：
undelete filename
S5700/6700/7700/9700 交换机：
undelete { filename | devicename }</td><td>恢复被删除到回收站中的文件（恢复时会有确认提示）。命令中的参数说明如下。
（1）filename：二选一参数，指定待恢复的文件名，其他说明同表 2-9 中介绍的 dir 命令的该参数说明
（2）devicename：二选一参数，指定要依次恢复指定存储器根目录下的所有被删除文件，取值可以是 flash:，cfcard:
当用户需要恢复之前删除过的文件或由于误操作删除某个文件时，只要不是永久删除（执行了带参数/unreserved 的 delete 命令或执行 reset recycle-bin 命令），都可以使用此命令将文件恢复。
恢复的文件名如果与同路径下现有的目录名重名，则执行失败；若与当前存在的文件名重名，将会提示是否覆盖</td></tr>
<tr><td>彻底删除
回收站
中的文件</td><td>S2700/3700 系列交换机：
reset recycle-bin [filename]
S5700/6700/7700/9700 交换机：
reset recycle-bin [filename |
devicename]</td><td>彻底删除指定路径下回收站中的文件，以释放空间。命令中的参数说明如下。
（1）filename：二选一可选参数，指定要彻底删除的文件名，其他说明同表 2-9 中介绍的 dir 命令的该参数说明
（2）devicename：二选一可选参数，指定要依次彻底删除指定存储器根目录下的所有回收站中的文件，取值可以是 flash:，cfcard:
如果不选择以上任何可选参数，则依次删除用户当前工作路径下回收站中的文件</td></tr>
</table>

（续表）

文件操作	所用命令	说明
执行指定的批处理文件	**execute** *batch-filename*	执行指定的批处理文件。当用户经常性的执行一系列命令时，则可以将这些命令逐条写入批处理文件，然后将此文件保存在交换机中，以后只需要执行此命令就可以完成之前手动输入执行的多条命令，帮助用户提升维护管理交换机的效率。 参数 *batch-filename* 为指定要执行的批处理文件名，以.bat 为后缀，但**必须在系统视图下执行**。批处理文件可以在文本编辑器中进行编辑，每一条需执行的命令占据一行，然后将文件扩展名".txt"替换为".bat"即可。编辑好的批处理文件需要通过文件传输方式上传至交换机中，具体上传方法将在下章介绍
配置文件系统提示方式	**file prompt** { **alert** | **quiet** }	修改文件操作的提醒方式。如果选择了 **alert** 二选一选项，则对用户进行的可能导致数据丢失或破坏的操作（比如删除文件操作等）需给用户确认和警告提示；如果选择了 **quiet** 二选一选项，则所有操作都不会有确认提示，直接执行 缺省情况下，为 **alert** 方式，建议不要修改，可用 **undo file prompt** 命令将文件操作提醒方式恢复为缺省的 **alert** 方式

【示例 1】显示当前目录下testcfg.cfg配置文件的内容。

```
< HUAWEI > more testcfg.cfg
#
sysname Sysname
#
configure-user count 5
#
vlan 2
#
return
<Sysname>
```

【示例 2】将文件config.cfg从cfcard存储器的根目录复制到cfcard:/temp目录中，目标文件名是temp.cfg。

```
<HUAWEI> copy cfcard:/config.cfg cfcard:/temp/temp.cfg
Copy cfcard:/config.cfg to cfcard:/temp/temp.cfg?[Y/N]:y
100% complete./
Info: Copied file cfcard:/config.cfg to cfcard:/temp/temp.cfg. .Done.
```

【示例 3】如果当前目录就是cfcard根目录，可以采用另一种方法（即采用相对路径方法）以完成上一个示例。

```
<HUAWEI>pwd
cfcard:
<HUAWEI>dir
Directory of cfcard:/
  Idx   Attr   Size(Byte)   Date   Time   FileName
  0   -rw-   6,721,804   Mar 19 2012 12:31:58   devicesoft.cc
  1   -rw-   910   Mar 19 2012 12:32:58   config.cfg
  2   drw-   -   Mar 05 2012 09:54:34   temp
..
509,256 KB total (52,752 KB free)
```

<HUAWEI> copy config.cfg temp/temp.cfg

Copy cfcard:/config.cfg to cfcard:/temp/temp.cfg?[Y/N]:y

100% complete./

Info: Copied file cfcard:/config.cfg to cfcard:/temp/temp.cfg. .Done.

【示例 4】将文件config.cfg从cfcard存储器的根目录复制到cfcard:/temp目录中，目标文件名与源文件名相同。

<HUAWEI>pwd

cfcard:

<HUAWEI>dir

Directory of cfcard:/

Idx Attr Size(Byte) Date Time FileName

0 -rw- 6,721,804 Mar 19 2012 12:31:58 devicesoft.cc

1 -rw- 910 Mar 19 2012 12:32:58 config.cfg

2 drw- - Mar 05 2012 09:54:34 temp

..

509,256 KB total (52,752 KB free)

<HUAWEI> copy config.cfg temp

Copy cfcard:/config.cfg to cfcard:/temp/config.cfg?[Y/N]:y

100% complete./

Info: Copied file cfcard:/config.cfg to cfcard:/temp/config.cfg. .Done.

【示例 5】当前工作路径是cfcard:/test/，将test目录中backup.zip文件备份保存到同目录的backup1.zip文件中。

<HUAWEI>pwd

cfcard:/test

<HUAWEI> copy backup.zip backup1.zip

Copy cfcard:/test/backup.zip to cfcard:/test/backup1.zip?[Y/N]:y

100% complete./

Info: Copied file cfcard:/test/backup.zip to cfcard:/test/backup1.zip. .Done.

【示例 6】把cfcard:/test/sample.txt文件移到cfcard:/sample.txt。

<HUAWEI>move cfcard:/test/sample.txt cfcard:/sample.txt

Move cfcard:/test/sample.txt to cfcard:/sample.txt ?[Y/N]: y

%Moved file cfcard:/test/sample.txt to cfcard:/sample.txt.

【示例 7】把flash:/test/sample.txt文件移动到flash:/sample.txt。

<HUAWEI>move flash:/test/sample.txt flash:/sample.txt

Move flash:/test/sample.txt to flash:/sample.txt ?[Y/N]: y

%Moved file flash:/test/sample.txt to flash:/sample.txt.

【示例 8】将cfcard:/test/路径下的mytest目录重命名为yourtest。

<HUAWEI>pwd

cfcard:/test

<HUAWEI> renamemytest yourtest

Rename cfcard:/test/mytest to cfcard:/test/yourtest ?[Y/N]:y

Info: Rename file cfcard:/test/mytest to cfcard:/test/yourtest . . .Done.

【示例 9】重命名文件sample.txt为sample.bak。

<HUAWEI> rename sample.txt sample.bak

Rename cfcard:/sample.txt to cfcard:/sample.bak ?[Y/N]:y

Info:Rename file cfcard:/sample.txt to cfcard:/sample.bak . . . .Done.

【示例 10】将根目录下log.txt文件压缩到test目录下log.zip文件。

<HUAWEI>dir

Directory of cfcard:/

Idx Attr Size(Byte) Date Time FileName

0 -rw- 155 Dec 02 2011 01:28:48 log.txt

1 -rw- 9,870 Oct 01 2011 00:22:46 patch.pat

2 drw- - Mar 22 2012 00:00:48 test

3 -rw- 836 Dec 22 2011 16:55:46 rr.dat

..

509,256 KB total (52,752 KB free)

<HUAWEI> zip log.txt cfcard:/test/log.zip

Compress cfcard:/log.txt to cfcard:/test/log.zip?[Y/N]:y

100% complete

%Compressed file cfcard:/log.txt to cfcard:/test/log.zip.

<HUAWEI> cd test

<HUAWEI>dir

Directory of cfcard:/test/

Idx Attr Size(Byte) Date Time FileName

0 -rw- 836 Mar 20 2012 19:49:14 test

1 -rw- 239 Mar 22 2012 20:57:38 test.txt

2 -rw- 1,056 Dec 02 2011 01:28:48 log.txt

3 -rw- 240 Mar 22 2012 21:23:46 log.zip

509,256 KB total (52,751 KB free)

【示例 11】当前工作路径是 cfcard:/selftest，要删除此路径下 cfcard:/selftest/test.txt文件。

<HUAWEI>delete test.txt

Delete cfcard:/selftest/test.txt?[Y/N]:y

Info: Deleting file cfcard:/selftest/test.txt. .succeeded.

【示例 12】恢复被删除的文件sample.bak。

<HUAWEI>undelete sample.bak

Undelete cfcard:/sample.bak ?[Y/N]:y

% Undeleted file cfcard:/sample.bak.

【示例 13】删除flash:根目录下test目录回收站中test.txt文件。

<HUAWEI> reset recycle-bin flash:/test/test.txt

Squeeze flash:/test/test.txt?[Y/N]:y

%Cleared file flash:/test/test.txt.

【示例 14】执行cfcard:/目录下的test.bat批处理文件。test.bat文件中包含以下命令：system-view、aaa、local-user huawei password cipher huawei@123。

<HUAWEI>system-view

[HUAWEI] execute test.bat

[HUAWEI]

^

Error: Unrecognized command found at '^' position.

[HUAWEI]

[HUAWEI-aaa]

Info: Add a new user

[HUAWEI-aaa]

[HUAWEI-aaa]

当执行第一条 system-view 命令时，因为当前正在系统视图下，所以提示错误，接着继续执行下面的命令，如添加了一个新的用户huawei。

2.3.4 存储器管理

在 PC 机中经常要进行磁盘的维护与管理，如格式化磁盘、修复文件系统，在网络交换机中同样需要类似的管理，那就是对它们的存储器进行维护和管理。在 S2700、S3700、S5700和S6700系列交换机中的存储器就是Flash:闪存，在S7700和S9700系列交换机中还有CF卡存储器。

1．格式化存储器

当文件系统的异常（如在dir命令的输出显示信息中含有unknown信息时）无法修复或者确认不再需要存储器上的所有数据时，可格式化存储器。但要注意，与格式化PC中的硬盘一样，格式化后会清空存储器中的所有文件和目录。

格式化存储器的方法与 DOS 下的格式化命令一样，也是 format，就是直接在用户视图下执行format devicename命令。这里的参数devicename就是指要格式化的存储器交换机名称，在S2700、S3700、S5700和S6700系列交换机只有flash:交换机，在S7700和S9700系列交换机中有主控板闪存“flash：”，备用主控板闪存“slave#flash：”，除此之外还有主控板CF卡“cfcard：”，备用主控板CF卡“slave#cfcard：”。

【示例 1】格式化flash：存储器。

<HUAWEI>format flash:

All data(include configuration and system startup file) on flash: will be lost, proceed with format ? [Y/N]: :y

%Format flash: completed.

2．修复文件系统

当存储器上的文件系统出现异常时，终端会给出提示信息，建议修复一下存储器上的文件系统。与PC机上的磁盘修复命令一样，VRP的文件系统修复命令也是fixdisk命令，其格式为fixdisk devicename，但不确保修复成功。命令中的参数devicename是指定要修复文件系统的存储器交换机，不同交换机上的存储器交换机同上面介绍。

【示例 2】终端显示如存储器CF卡出错，进行修复。

Lost chains in cfcard detected, please use fixdisk to recover them!

<HUAWEI> fixdisk cfcard:

% Fix disk cfcard: completed.

2.4 VRP系统的组成

VRP系统在启动时需要加载“系统软件”和“配置文件”两部分，这与其他品牌网络交换机的操作系统是一样的。如果指定了下次启动的补丁文件，还需加载补丁文件。修改VRP系统启动的场景一般有以下几种。

1．对交换机进行升级操作，即系统软件从低版本升级至高版本

当增加了新特性或者需要对原有性能进行优化以及解决当前运行版本落后的问题时，则需要对交换机进行升级。此时需要加载高版本的系统软件，并重新启动交换机来实现。

2．对交换机进行降级操作（版本回退），即系统软件从高版本降级至低版本

交换机完成升级后，如果业务出现异常，为保证业务正常可以先将交换机版本进行回退。此时需要加载低版本的系统软件，并重新启动交换机来实现。

3．对一个新交换机加载已有的满足用户需求的配置文件

新交换机中只包含了出厂时的缺省配置，如果需要使这台新交换机连接至网络再运行业务，则需要用户在交换机上进行大量的配置，花费不少时间。对于这种情况，只需要为这台新交换机指定满足用户需求的配置文件，然后重新启动交换机即可，大大提升了用户对交换机的配置效率。

4．对交换机指定升级后的补丁文件

可在交换机升级的同时指定之前未安装过的补丁文件，升级完成后补丁也会立即生效。

2.4.1 VRP系统软件

华为VRP系统包括“软件系统”和“配置文件”两大部分，本节先介绍VRP软件系统，下节将介绍VRP配置文件。

华为S系列交换机的VRP软件系统包括“BootROM软件”和“系统软件”两部分，分别如PC机主板芯片上固化的BIOS系统和硬盘中安装的各种操作系统。交换机加电后，先运行BootROM软件，初始化硬件并显示交换机的硬件参数，再运行系统软件。系统软件一方面提供对硬件的驱动和适配功能，另一方面实现了业务功能特性；BootROM 软件与系统软件是交换机启动、运行的必备软件，为整个交换机提供支撑、管理、业务等功能。

交换机在升级时包括升级 BootROM 软件和升级系统软件。目前交换机的系统软件中已经包含了BootROM软件，所以在升级系统软件的同时即可自动升级BootROM软件。也正因如此，现在所说的VRP系统软件其实就代表了整个VRP软件系统。

1．VRP系统软件版本

华为VRP系统软件版本分为“核心版本”（或者“内核版本”）和“发行版本”两种。其中的核心版本是用来开发具体交换机VRP系统的基础版本，也就是通常所说的VRP 1.x、2.x、3.x，以及现在的VRP 5.x和8.x版本；发行版本则是在核心版本基础上针对具体的产品系列（如有S系列交换机系列、AR/NE系列路由器系列等）而发布的VRP系统版本。

VRP系统的核心版本是由一个小数来表示，小数点前面的数字表示主版本号，仅当发生比较全面的功能或者体系结构修改时才会发布新的主版本号；小数点后面第1位数字表示次版本号，仅当发生重大或者较多功能修改时才会发布新的次版本号；后面1～2位数字为修订版本号，只要发生修改就会发布新的修订版本号。如上面的VRP 5.120中的主版本号为5，次版本号为1，20为修订版本号。

华为VRP系统的发行版本是以V、R、C三个字母（代表三种不同的版本号）进行标识的，基本格式为

VxxxRxxxCxx，其中的x是一些具体的数字。V、R部分为必须部分；C根据版本性质的不同而确定，可能出现也可能不出现。V、R、C这三个字母的定义如下。

（1）V版本是指产品所基于的软件或者硬件平台版本。

Vxxx标识产品/解决方案主力产品平台版本的变化，称为V版本号。其中的xxx从100开始，并以100为单位递增编号。仅当产品的平台发生变化，V版本号才会发生变化。

（2）R版本是面向客户发布的通用特性集合，是产品在特定时间的具体体现形式。

Rxxx标识面向所有客户发布的通用版本，称为R版本号。其中的xxx从001开始以1为单位递增编号。

注意

上述V版本号和R版本号独立编号，互不影响。也就是它们之间并没有从属关系。 例如产品平台发生变化，而功能特性不变，如原VR版本号为V100R005，则新的VR版本号为 V200R005。当然，若产品功能特性发生变化，平台却不变。根据这一原则可以得出，基于 V100R005 升级的后一个版本的版本号只可能是V100R006、V200R005、V200R006中的任意一种。

（3）C版本是基于R版本开发的快速满足不同类型客户需求的客户化版本。

在同一R版本下，C版本号中的xx从00开始以1为单位递增编号。如果R版本号发生变化，C版本号下的xx又从01开始重新编号，如V100R001C01、V100R001C02、V100R002C01。

以上这两个 VRP 系统版本均可通过 display version 命令查到。下面是一个执行display version命令的输出示例，其中的Version5.120就代表当前交换机运行的VRP核心版本为 5.120，而括号里面的“S5700 V200R002C00”则是指S5700系列交换机的VRP发行版本。同样还可从中看到对应的BootROM软件版本，如其中的“Basic BOOTROM Version : 100”表示BootROM软件版本号为 100。当然还可查看许多其他版本信息，如PCB印制电路板版本（Pcb Version）、复杂可编程逻辑交换机版本（CPLD Version，也即可编程芯片的版本）等。

```
<HUAWEI>display version
Huawei Versatile Routing Platform Software
VRP (R) software, Version 5.120 (S5700 V200R002C00)
Copyright (C) 2000-2012 HUAWEI TECH CO., LTD
HUAWEI S5700-52C-EI Routing Switch uptime is 0 week, 2 days, 1 hour, 24 minutes
EMGE 0(Master) : uptime is 0 week, 2 days, 1 hour, 23 minutes
512M bytes DDR Memory
64M bytes FLASH
Pcb    Version : VER B
Basic BOOTROM Version : 100 Compiled at Mar 1 2011, 20:27:16
CPLD    Version : 74
Software Version : VRP (R) Software, Version 5.120 (S5700 V200R002C00)
FANCARD information
Pcb    Version : FAN VER B
PWRCARD I information
Pcb    Version : PWR VER A
```

2．VRP系统软件名称

我们一般所说的系统软件是指产品版本的 VRP 系统软件。VRP 系统软件的文件扩展名为“.CC”，如V200R002C00.CC，如果要针对特定子系列，则在前面还会加子系列名，如 S5700HI-V200R002C00.CC。但

在华为公司网站下载的文件是.zip 格式的压缩文件，要解压后才能上传到交换机存储器中使用。

2.4.2 VRP系统配置文件

VRP 系统配置文件是 VRP 命令行的集合，用户可将当前配置保存到配置文件中， 以便在交换机重启后这些配置能够继续生效。另外，通过配置文件用户可以非常方便地查阅配置信息，也可以将配置文件上传到其他的交换机上，实现交换机的批量配置。

配置文件为文本文件，其规则如下。

（1）以命令格式保存。

（2）为了节省空间，只保存非缺省的参数。

（3）以命令视图为基本框架，同一命令视图的命令组织在一起，形成一节，节与节之间通常用空行或注释行隔开（以“#”开始的为注释行）。空行或注释行可以是一行或多行。

（4）文件中各节的顺序安排通常为全局配置、接口配置、各种协议配置和用户界面配置。

（5）配置文件必须以“.cfg”或“.zip”作为扩展名，而且必须存放在存储交换机的根目录下。

交换机在运行过程中，有配置文件和当前配置，它们的区别如表2-11所示。

表2-11 配置文件和当前配置的区别

配置文件类型	说明	查看方式
配置文件	交换机上电时，从缺省存储路径中读取配置文件进行交换机的初始化操作，因此该配置文件中的配置称为初始配置。如果缺省存储路径中没有配置文件，则交换机用缺省参数初始化配置	使用 **display startup** 命令可以查看到交换机本次以及下次启动的配置文件 使用 **display saved-configuration** 命令可以查看交换机下次启动时的配置文件信息
当前配置	与初始配置相对应，交换机运行过程中正在生效的配置称为当前配置，可以与配置文件的内容不一致，当然也可能是一致的，如当前没有做任何配置修改时	使用 **display current-configuration** 命令查看交换机的当前配置信息

用户通过命令行接口可以修改交换机当前配置，为了使当前配置能够作为交换机下次启动时的起始配置，需要使用 save 命令保存当前配置到缺省存储器中，形成配置文件。

说明

配置文件支持包含 30 000条命令行。如果超过了 30 000条，在交换机进行升级时， 不能保证所有命令在升级后兼容。

如果使用不完整格式进行配置，由于命令保存到配置文件中时使用的是完整格式，可能导致配置文件中存在长度超过510个字符的命令（系统可正确执行的命令长度最大为510个字符）。系统重启时，这类命令将无法恢复。

2.4.3 VRP系统补丁文件

补丁是一种与交换机VRP系统软件兼容的软件，用于解决交换机系统软件少量且急需解决的问题，就像各种操作系统（如 Windows 系统）、应用软件陆续发布的补丁文件一样。在交换机的运行过程中，有时需要对交换机系统软件进行一些适应性和排错性的修改，如改正系统中存在的缺陷、优化某功能以适应业务需求等。

补丁通常以补丁文件的形式发布，一个补丁文件可能包含一个或多个补丁，不同的补丁具有不同的功能。当补丁文件被用户从存储器加载到内存补丁区中时，补丁文件中 的补丁将被分配一个在此内存补丁区中唯一的单元序号，用于标志、管理和操作各补丁。

1．按补丁的适用范围分类及补丁编号分类

补丁文件分为产品补丁（适用于某个特定的VRC版本VRP系统）和公共补丁（适用于所有使用相同VR版本VRP系统的交换机），都有一个对应的补丁编号。

（1）产品补丁仅适用于对特定交换机的补丁软件，其编号是在特定交换机的 VRC版本的最后面再加上 SPCXXX，其中的 XXX 是代表补丁编号的 3 位数字，如V200R001C00SPC300中最后的SPC300就代表补丁编号为300。

（2）公共补丁是可适用于某个VR版本的VRP系统的通用补丁，其编号是在VR版本的最后面加上SPHXXX，其中的 XXX 表示公共补丁编号的 3 位数字，如V200R001SPH002中最后的SPH002就代表补丁编号为002。

2．按补丁生效对业务的影响分类

根据补丁生效对业务运行的影响分成热补丁和冷补丁。

（1）热补丁HP（Hot Patch）：补丁生效不中断业务，不影响业务运行，同时可以降低交换机升级成本，避免升级风险。

（2）冷补丁CP（Cold Patch）：要使补丁生效需要复位单板或重启交换机，影响业务的运行。

3．按补丁间的依赖性分类

根据补丁间的依赖关系，补丁可分为增量型补丁和非增量型补丁。

（1）增量型补丁：是指对在其前面的补丁有依赖性的补丁。一个新的补丁文件必须包含前一个补丁文件中的所有补丁信息。用户可以在不卸载原补丁文件的情况下直接安装新的补丁文件。

（2）非增量型补丁：只允许当前系统安装一个补丁文件。如果用户安装完补丁之后希望重新安装另一个补丁文件，则需要先卸载当前的补丁文件，再重新安装并运行新的补丁文件。

目前，产品发布的补丁类型都为热补丁与增量型补丁。在后续的描述中如无特别说明都是指此类补丁。

4．补丁状态

每个补丁都有自身的状态，只有在用户命令行的干预下才能发生切换。补丁状态详细信息如表2-12所示。

表2-12 补丁状态

状态	说明	各状态之间的转换关系
空闲态（Idle）	此时，补丁文件存储在交换机的存储器中，但文件中的补丁还没有被加载到内存补丁区中	当用户将补丁从存储器中加载到内存补丁区后，补丁的状态将被设置为去激活
去激活（Deactive）	当补丁被加载到内存补丁区中或激活的补丁被停止运行时，补丁就处于去激活状态	用户可以对去激活状态的补丁进行以下两种操作。 • 卸载此补丁，使补丁从内存补丁区中被删除 • 临时运行此补丁，使补丁的状态变为激活状态

（续表）

状态	说明	各状态之间的转换关系
激活（Active）	当补丁被存储在内存补丁区中，且被临时运行时，补丁就处于激活状态 当单板被复位后，此单板上在复位前处于激活状态的补丁仍然恢复为激活状态。只有当整机复位后，复位前处于激活状态的补丁将会处于去激活状态	用户可以对激活状态的补丁进行以下 3 种操作。 • 卸载此补丁，使补丁从内存补丁区中被删除。 • 停止运行此补丁，使补丁的状态变为去激活状态 • 永久运行此补丁，使补丁的状态变为运行状态
运行（Running）	当补丁被存储在内存补丁区中，且被永久运行时，补丁就处于运行状态 当单板或整机被复位后，在复位前处于运行状态的补丁将保持运行状态	用户可以卸载处于运行状态的补丁，使补丁从内存补丁区中被删除

2.4.4 启动BootROM软件

华为 S 系列交换机的 VRP 软件系统包括 BootROM 软件和系统软件两部分，其中BootROM软件又分为基本BootROM软件和扩展BootLoad软件。交换机上电后，先运行基本BootROM软件，并负责引导运行BootLoad软件，BootLoad软件负责引导运行系统软件。注意，不同版本的 BootROM 软件，下面的运行提示信息可能有较大区别。下面仅以S7700系列为例进行介绍。

1．基本BootROM软件的启动过程

在交换机上电后，首先运行基本 BootROM 软件，交换机的硬件开始自检，显示信息如下：

input 'm' to Select Debug Console:

Boardname . . . . .SRU

L2 Cache Test Start ? ('t' or 'T' is test). . . . . .OK

BIOS Creation Date . . . . Mar 9 2010, 22:34:36

Bootbus init. . . . . .OK

DDR DRAM init. . . . . .OK

Start Memory Test ? ('t' or 'T' is test):skip

Copying Uncompressed Data from Rom to Ram . . . . .Done

Uncompressing Data from Rom to RAM .Done

Initializing Flash Module . . . . .Done

如果在上述"L2 Cache Test Start? ('t' is test)"提示信息处按下T键（代表要进行测试）则进行二级Cache的检测，否则跳过（skip）；如果在"Start Memory Test ? ('t' or 'T' is test)"提示信息处按下T键则进行内存的检测，否则跳过（skip）。此时，如果需要检测内存，请在2s内按下<Ctrl+T>组合键。屏幕显示以下信息。

Testing DDR SDRAM, please wait for a few minutes

The detected DDR SDRAM size is: 1024MB

Testing DDR SDRAM: 1024MB . . . . pass

Took time: 23s

当屏幕显示以下信息时，如果在2s内按下<Ctrl+A>组合键，则进入基本BootROM菜单，否则继续后面的BootLoad软件启动过程，即执行基本BootROM菜单中的第4项。

Press Ctrl+A to enter Bootrom Menu. .

基本BootROM菜单界面如下（#后面是加和注释）。

Update Bootrom Menu (Ver 102)

Creation date: Mar 6 2009, 15:59:02

1. Update bootrom through serial interface　#---通过串口更新基本BootRom软件

2. Update bootload through serial interface #---通过串口更新扩展BootLoad软件

3. Modify serial interface parameter #---编辑串口参数

4. Boot from bootload system #---从扩展BootLoad软件启动系统

5. Reboot #---重启系统

通过基本BootROM菜单，可以升级基本BootROM软件（第1项）、扩展BootLoad软件（第2项）和编辑串口参数（第3项）。

2．BootLoad软件的启动过程

当没有在前面按下<Ctrl+A>组合键，或者在以上BootROM菜单中选择执行第4项，即启动BootLoad软件，开始初始化硬件并显示交换机的硬件参数信息。显示信息如下：

****

* *

* Ethernet Switch Bootload, Ver 121 *

* *

****

Copyright(C) 2003-2011 by HUAWEI TECHNOLOGIES CO., LTD.

Creation date: Apr 18 2012, 11:15:46

PCB Version : LE02SRUA VER.D

CPU L2 Cache : 128KB

CPU Clock Speed : 700MHz

BUS Clock Speed : 133MHz

Memory Type : DDR2 SDRAM

Memory Size : 1024MB

Memory Speed : 667MHz

CF Card Init. . . . . .Done

Description data is vaild in Nvram area !

Press Ctrl+B to enter Boot Menu. . 0

如果用户在3s内按下<Ctrl+B>组合键，则提示用户输入进入扩展BootROM菜单的密码（缺省密码是Admin@huawei.com），屏幕显示如下信息。

password:

此密码在系统视图下可通过reset boot password重置为缺省密码Admin@huawei.com。

输入正确的密码后，则进入BootLoad菜单。BootLoad菜单的界面如下。

1. Boot with default mode #---使用缺省模式启动系统

2. Boot from Flash #---从flash:闪存启动系统

3. Boot from CFCard #---从CF卡启动系统

4. Enter serial submenu #---进入串口子菜单

5. Enter ethernet submenu #---进入以太网口子菜单

6. Modify Flash description area #---修改闪存描述区域

7. Modify bootrom password #---修改进入基本BootRom菜单的密码

8. Clear password for console user #---清除控制台用户密码（在用户忘记密码时可用）

9. Reboot #---重启系统

Enter your choice(1-9):1

通过 BootLoad 菜单，用户指定交换机启动时加载的系统软件，修改进入基本BootRom 菜单密码，清除Console 用户密码等。缺省执行第 1 项菜单，初始化串口和Console，解压缩系统软件，并引导运行系统软件。屏幕显示的信息如下。

Auto-booting. .

Booting from CFCard. .

Loading.............................Done!

Uncompressing. .Done!

至此，BootROM软件引导过程结束，交换机将开始加载系统软件。

2.5 管理VRP配置文件

前面说了VRP系统有"配置文件"（已以文件形式保存的配置）和"当前配置"（正在运行、生效的配置，仅指没有以文件形式保存的配置）两种配置文件。用户可以进行保存配置文件（即把当前配置以文件形式保存起来）、备份配置文件（备份已有的配置文件）、恢复配置文件（恢复使用其他配置文件），指定下次启动的启动文件（包括配置文件）等操作。下面分别予以介绍。

2.5.1 保存配置文件

用户可以通过命令行修改交换机的当前配置，而这些配置在设备重启后将失效；如果要使当前配置在系统下次重启时仍然有效，在重启交换机前需要将当前配置保存到配置文件中。可以采用"自动保存配置"和"手动保存配置"两种方法保存配置文件。

1．自动保存配置文件

自动保存配置文件分两种情况：一种是自动保存配置文件在本地交换机存储器中，另一种自动保存配置文件在远程的服务器上。

（1）本地自动保存配置文件

在本地自动保存配置文件的方法是需要先在系统视图下使用 set save-configuration [interval interval | cpu-limit cpu-usage |delaydelay-interval] *命令配置系统定时保存配置文件。命令中的参数说明如下。

① interval：可多选参数，指定定时保存配置的时间间隔（即每隔这个时间自动保存一次配置文件），取值范围为（30～43 200）整数分钟。缺省值是 30min。

② cpu-usage：可多选参数，指定定时自动保存时的CPU占用率阈值（高于这个阈值即取消当前进行的自动进行配置文件保存操作），取值范围是1～60的整数值，代表对应的百分比，缺省值为50%。这是为了防止自动保存影响系统性能。

③ delay-interval：可多选参数，指定配置变更发生后系统自动备份配置的延时时间（即在发生配置文件更改后多少时间自动进行配置文件保存），取值范围为（1～60）整数分钟，但其取值必须小于同时设置的interval参数值。缺省值是5min。

配置系统定时自动保存功能后，会把配置文件保存在下次启动配置文件中，配置文件内容可能会因配置变化而变化。如果没有配置本命令，则系统也不启动自动保存功能。 但是系统在定时保存配置之前会比较配置文件，如果配置没有改变则不会执行定时保存，即使符合了本命令设置的各参数值条件。

缺省情况下，VRP系统不启动定时保存配置的功能，可用undo set save-configuration [interval interval | cpu-limit cpu-usage | delay delay-interval] *命令取消原来的自动配置文件保存设置。当出现如下情况时，系

统会取消定时保存配置文件的操作。

① 当前存在写配置文件操作。

② 接口板正在进行配置恢复。

③ CPU利用率较高。

【示例 1】设置系统定时保存新配置的时间间隔为60min。

<HUAWEI> system-view

[HUAWEI] set save-configuration interval 60

【示例 2】设置在系统配置发生变化3min后，以10h为保存间隔，自动保存新配置，且CPU使用率上限为60%。

<HUAWEI> system-view

[HUAWEI] set save-configuration interval 600 delay 3 cpu-limit 60

（2）远程保存配置文件

如果要把配置文件自动保存在远程服务器上，则需要先通过 set save-configuration backup-to-server server server-ip transport-type { ftp | sftp }user user-name password password [path folder] 或 set save-configuration backup-to-server server server-ip transport-type tftp [path folder] 系统视图命令分别配置FTP、SFTP或TFTP服务器的相关信息，包括自动保存配置文件的对应服务器的IP地址、用户名及其密码、配置文件自动保存的目的路径，采用FTP、SFTP或者TFTP对应的传输方式把配置文件自动保存至对应的服务器上（需要事先在对应的终端 PC 上配置好对应的服务器，并确保交换机与服务器之间的路由可达）。命令中的参数和选项说明如下。

① server-ip：指定定时保存配置文件的FTP、SFTP或者TFTP服务器IP地址。

② ftp ：二选一选项，指定采用FTP作为文件传输协议，把配置文件自动保存到指定的FTP服务器上。

③ sftp ：二选一选项，指定采用SFTP作为文件传输协议，把配置文件自动保存到指定的SFTP服务器上。

④user-name：指定访问FTP或者SFTP服务器的用户名（TFTP服务器访问不需要配置用户名和密码，因为它是采用UDP传输层协议进行通信的 ），为1～64个字符，不支持空格，区分大小写。

⑤password：指定访问服务器的用户密码，明文密码为1～16个字符，密文密码为32个字符，不支持空格，区分大小写。

⑥ folder：可选参数，指定服务器存储配置文件的相对路径，为1～64个字符，不支持空格，区分大小写。如果不指定此可选参数，则自动把配置文件保存在当前服务器所在目录下。如果指定的路径不存在，则配置文件将发送不成功，系统将向网管上报告警，并在交换机上记录日志。

【示例 3】把配置文件自动以用户名为huawei，密码为huawei2012保存到IP地址为1.1.1.1的SFTP服务器上。

<HUAWEI> system-view

[HUAWEI] set save-configuration backup-to-server server 1.1.1.1 transport-type sftp user huawei password huawei2012

说明

使用TFTP传输方式保存配置文件时，可使用 tftp client-source{ -a source-ip-address | -i interface-type interface-number }命令配置交换机的Lookback接口和其 IP地址作为当交换机作为TFTP客户端发送报文的源接口和源IP地址。缺省情况下，TFTP客户端发送报文的源地址为0.0.0.0。

2．手动保存配置文件

如果你没有配置以上的自动保存配置文件，或者因为刚发生的配置更改很重要，你想立即保存，则可进行手动保存配置文件。手动保存仅会保存在交换机本地存储器中，方法是执行 save [all] [configuration-file] 配置，保存当前配置。命令中的参数和选项说明如下。

（1）all：可选项，选择它后将保存所有的配置，包括不在位的板卡的配置。

（2）configuration-file：可选参数，指定所保存的配置文件名称（包括路径），绝对路径的长度范围为5～64 个字符。在第一次保存配置文件时，如果不指定可选参数configuration-file，则交换机将提示是否将文件名保存为“vrpcfg.zip”。“vrpcfg.zip”是系统缺省的配置文件，初始状态是空配置。

将当前配置保存到指定文件时，文件必须以“.zip”或“.cfg”作为扩展名，而且系统启动配置文件必须存放在存储交换机的根目录下 。*.cfg为纯文本格式，可直接查看里面的内容，指定为配置文件后，启动时系统对里面的命令逐条进行恢复；*.zip 是*.cfg的压缩，占用空间较小，指定为配置文件后，启动时要先解压成*.cfg格式，然后逐条恢复。注意以下几种命令格式的作用效果。

（1）执行不带任何参数和选项的save 命令将直接替换当前启动配置文件中的相应内容。多数情况下是这样直接保存的。

（2）执行 save all 命令会保存当前所有的配置到当前启动配置文件中（直接替换相应内容），包括不在位的板卡配置。

（3）执行save configuration-file 命令将保存当前配置信息到交换机中指定的配置文件中。通常情况下不影响系统当前的启动配置文件，除非当 configuration-file 与系统缺省的存储路径及配置文件名完全相同时，此时就等同于save 命令。

（4）执行save all configuration-file 命令用来保存当前配置信息到交换机中指定的配置文件中。通常情况下也不影响系统当前的启动配置文件，除非当 configuration-file与系统缺省的存储路径及配置文件名完全相同时，此时就等同于save all 命令。

【示例 4】使用save 命令直接保存当前配置文件到缺省存储交换机中。

```
<HUAWEI>save
The current configuration will be written to the device.
Are you sure to continue?[Y/N]y
Now saving the current configuration to the slot 0.
Save the configuration successfully.
```

2.5.2 备份配置文件

为防止交换机或者配置文件意外损坏而导致配置文件无法恢复，可以通过以下4种 方法进行配置文件备份。

- 直接屏幕复制。
- 备份配置文件到存储器其他位置。
- 通过TFTP备份配置文件到远程TFTP服务器中。
- 通过FTP备份配置文件到远程FTP服务器中。

1．直接屏幕复制

直接屏幕复制的方法是最原始的方式，可先在命令行界面上，执行 display current-configuration 命令并复制所有显示信息到TXT文本文件中，从而将配置文件备份到维护终端的硬盘中。注意配置文件的扩展名一定要为.cfg。

2．备份配置文件到flash:或cfcard:存储器中

可以把配置文件以非缺省配置文件名备份保存在交换机当前的 flash：或者 cfcard:存储器中。在交换机启动之后，使用copy命令备份配置文件。下面是一个示例，把当前配置文件config.cfg以配置文件名backup.cfg备份到存储器根目录下。

<HUAWEI>save config.cfg

<HUAWEI> copy config.cfg backup.cfg

如果不是保存在交换机的缺省存储器根目录下，需要指定绝对路径。

3．通过TFTP备份配置文件

这种备份方式是将当前交换机作为 TFTP 客户端，然后在通过网络相连的 PC 机上配置并启动TFTP服务器程序。设置好下载配置文件的传输路径、TFTP服务器IP地址、端口号。然后在本地交换机用户视图下执行 tftp [-a source-ip-address | -i interface-type interface-number] tftp-server put source-filename [destination-filename] 命令备份指定的配置文件。命令中的参数说明如下。

（1）source-ip-address：二选一可选参数，指定本端交换机的 IP 地址，用户可以以指定的IP地址与服务端通信，从而达到进行安全校验的目的。

（2）interface-type interface-number：二选一可选参数，指定本端交换机出接口的接口类型和接口编号。

（3）tftp-server：指定TFTP服务器的IP地址或者主机名。

（4）source-filename：指定备份的源配置文件名。

（5）destination-filename：可选参数，指定备份后的目标配置文件名。如果不指定此可选参数，则与源配置文件名一样。

下面是一个配置文件备份示例，把 cfcard:存储器中的配置文件以备份配置文件名backup.cfg保存到IP地址为10.110.24.254的TFTP服务器根目录下。

<HUAWEI> tftp 10.110.24.254 put cfcard:/config.cfg backup.cfg

4．通过FTP备份配置文件

通过FTP服务备份配置文件的方式有两种：一种是把当前交换机作为FTP服务器，把用来保存备份配置文件的PC机作为FTP客户端。通过PC机提示符下键入的ftp命令向交换机发起FTP连接，然后在PC机上通过FTP服务的get命令从交换机上备份配置文件。这种方法相对来说比较简单，因为在交换机上启用 FTP 服务器功能比较容易。

另一种则相反，把当前交换机作为 FTP 客户端，而把用来保存备份配置文件的 PC机作为FTP服务器。通过交换机命令行中键入的ftp命令向PC机发起FTP连接，然后在交换机上通过FTP服务的put命令把交换机上的配置文件备份到PC机上。这种方法需要先在PC机上安装并配置好FTP服务器，相对来说比较麻烦。

下面仅介绍第一种方法，需要经过以下几个步骤（详细配置任务可参见本书第3章3.7.2节）：

（1）在当前交换机上启动FTP服务器功能，并创建用户名（假设为huawei）和密码（假设为huawei@123）的FTP用户，授权此用户可访问配置文件保存的存储器（可以是flash:或cfcard:）。下面是访问cfcard:存储器的配置示例。

<HUAWEI>system-view

[HUAWEI] ftp server enable　#---启用FTP服务器功能

Info: Succeeded in starting the FTP server.

[HUAWEI] aaa　#---启用AAA认证，进入AAA视图

[HUAWEI-aaa] local-userhuaweipassword cipher huawei@123　#---配置本地用户huawei，加密密码为huawei@123

[HUAWEI-aaa] local-userhuawei ftp-directory cfcard: #---配置用户huawei可访问的目录为 cfcard:存储器

[HUAWEI-aaa] local-userhuawei service-type ftp   #---配置用户huawei可以使用FTP服务

[HUAWEI-aaa] local-user huawei privilege level 15   #---配置用户huawei的用户级别为最高的15级

（2）在PC上向交换机发起FTP连接（假设交换机的管理IP地址是10.110.24.254），以建立与交换机的FTP连接。

C:\Documents and Setting\Administrator> ftp 10.110.24.254

Connected to 10.110.24.254.

220 FTP service ready.

User (10.110.24.254:(none)):huawei

331 Password required for huawei.

Password:

230 User logged in.

（3）设置传输参数。FTP用户验证通过后，在FTP客户端会显示 “ftp>”提示符，键入binary（二进制传输模式），并设置FTP客户端存放上载文件的目录路径（假设为C:\temp）。

ftp>binary

200 Type set to I.

ftp> lcd c:\temp

Local directory now C:\temp.

（4）传输配置文件。在FTP客户端PC上，使用get命令将配置文件下载至本地指定目录中，并保存为backup.cfg。如果文件大小一致则认为备份成功。

ftp>get cfcard:/config.cfg backup.cfg

2.5.3 恢复配置文件

如果用户进行了错误的配置，或者原来的配置文件已损坏，将导致交换机某些功能异常，此时可以通过以下3种方法进行配置文件恢复。

（1）从存储器上备份的配置文件中恢复配置文件。

（2）通过TFTP恢复备份在PC上的配置文件。

（3）通过FTP恢复备份在PC上的配置文件。

在恢复配置文件后，为了让配置文件生效需要重新启动交换机。先使用 startup saved-configuration configuration-file命令指定重新启动使用的配置文件（如果配置文件命名没有变，则该步骤省略），然后使用reboot命令重新启动交换机。

1．从存储器恢复配置文件

这种恢复方法主要便于用户将存储在交换机存储器（可以是flash:或cfcard:）中的备份配置文件恢复成当前系统运行的配置文件。在交换机正常工作时，可使用如下命令恢复配置文件（假设原来的备份配置文件名是保存在 cfcard:存储器根目录下的backup.cfg）。

<HUAWEI> copy cfcard:/backup.cfg cfcard:/config.cfg

然后通过 startup saved-configuration configuration-file命令指定复制的配置文件为下次启动时所用的配置文件。该命令具体将在本章2.6.1节介绍，在此不再赘述。

2．通过TFTP恢复备份在PC上的配置文件

这种恢复方法是将当前交换机作为TFTP客户端。恢复的方法与上节介绍的“通过TFTP备份配置文件”中

的备份步骤相同，差别仅在于在命令行界面中执行携带get参数的tftp命令，将存储在PC上的配置文件backup.cfg下载到交换机的cfcard中。

3．通过FTP恢复备份在PC上的配置文件

这里同样可以有两种方法，参见上节“通过FTP备份配置文件”中的介绍。当把当前交换机作为FTP服务器时恢复备份配置文件的步骤与上节“通过FTP备份配置文件”中的步骤相同，差别仅在于在最后传输配置文件时需在命令行界面中执行put命令，将存储在PC上的配置文件backup.cfg上传到交换机的cfcard中。

2.5.4 比较配置文件

设备使用的时间一长，里面存放的配置文件可能就比较多，容易造成混乱（通常可通过为配置文件起一个有隐含意义的文件名来进行区分）。这时就可以把某个配置文件与当前配置进行比较，通过配置结果中显示的配置不同处比较出各个配置文件的版本新、旧，从而决定可以删除某些不再使用的配置文件，也可以决定是否需要将当前配置设置为下次启动时加载的配置文件。

通过配置文件的比较，VRP系统在比较出不同之处时将从两者有差异的地方开始显示字符，（缺省显示150个字符），如果该不同之处到文件末尾不足150个字符，将显示到文件尾为止。所比较的配置文件必须以“.cfg”或“.zip”作为扩展名。如果指定要与当前配置进行比较的配置文件不存在，或者虽然配置文件存在，但是内容为空，系统将提示读文件失败。

在系统视图下执行 compare configuration [configuration-file] [current-line-number save-line-number] 命令，可以比较当前配置与指定的配置文件或者指定的配置文件的内容是否一致。命令中的参数说明如下。

（1）configuration-file：可选参数，指定需要与当前配置进行比较的配置文件名，长度范围为5～48个字符，不支持空格。如果不指定此可选参数，系统将比较当前的配置与下次启动配置文件内容是否一致。

（2）current-line-number save-line-number：可选参数，指定在当前配置中从指定的行开始比较，在指定配置文件中从指定的行开始比较。如果不指定此可选参数，则表示从指定的配置文件的首行开始进行比较。用来指定在发现配置文件不同之处后，跳过该不同处各自从指定的行继续进行比较。

【示例 】比较当前配置与下次启动的配置文件内容是否一致。从输出信息可以看出，这两个配置文件中均从第6行开始不一致，并且分别列出了两个配置文件中的对应配置。

```
<HUAWEI> compare configuration
Warning: The current configuration is not the same as the next startup configura
tion file.
====== Current configuration line 6 ======
vlan batch 1 to 2 10 to 11 15 70 to 71 91 to 92 100 111 230 240 901
vlan batch 911 1111
#
l2protocol-tunnel vtp group-mac 0100-0ccd-ffff
====== Configuration file line 6 ======
vlan batch 1 to 2 10 to 11 15 70 91 to 92 100 111 230 240 901
vlan batch 911 1111
#
l2protocol-tunnel vtp group-mac 0100-0ccd-ffff
```

2.5.5 清除配置文件

在以下情况下需要清除（删除）配置文件。

（1）交换机软件升级之后，原配置文件与当前软件不匹配。

（2）配置文件遭到破坏，或加载了错误的配置文件。

可在系统视图下执行 reset saved-configuration命令，清除当前加载的配置文件。系统在清除交换机配置文件前会比较当前启动与下次启动的配置文件。

（1）如果一致，执行该命令将同时清除这两个配置文件。此时可以在交换机上设定下次启动文件，否则下次启动时配置文件为空。

（2）如果不一致，执行该命令将清除当前启动的配置文件。

（3）如果交换机当前启动的配置文件为空，执行该命令后，系统将提示配置文件不存在。

执行该命令后，如果不使用 startup saved-configuration configuration-file命令重新指定含有正确配置信息的配置文件，或者不使用save命令保存配置文件，则交换机下次启动时将采用缺省的配置参数进行初始化。

如果是S5700、S6700、S7700、S9300和S9700系列交换机，则还可一键式清除指定接口下配置信息或将其配置恢复到缺省值。只需在系统视图下执行 clear configuration interface interface-type interface-number命令，即可清除指定接口下配置信息或将其配置恢复到缺省值；也可在对应接口视图下执行 clear configuration this命令清除该接口（但不支持Tunnel和css-port类型接口）下配置信息或将其配置恢复到缺省值。被清除配置文件的接口将被置为shutdown状态。

2.6 交换机启动管理

交换机启动管理包括指定系统启动文件和重启操作。配置系统启动文件包括指定系统启动时所用的系统软件和配置文件，这样可以保证交换机在下一次启动时以指定的系 统软件启动和指定的配置文件初始化配置。如果系统启动时还需要加载新的补丁，则还需指定补丁文件。但所指定的启动文件必须已保存至交换机的根目录中。

2.6.1 配置系统启动文件

系统启动文件也是在用户视图下配置 的。在进行系统启动文件配置前，可使用display startup 命令查看当前交换机指定的下次启动时加载的文件。如果没有重新配置交换机下次启动时加载的系统软件，则下次启动时将缺省使用本次加载的系统软件。当需要更改下次启动的系统文件（如交换机升级）时，则需要重新指定下次启动时加载的系统软件，此时还需要提前将系统软件通过文件传输方式保存至交换机上（系统软件必须存放在存储器的根目录下，且文件扩展名必须为“.cc” ）；如果交换机是双主控环境，需要确保系统软件分别保存至主用主控板和备用主控板存储器上。

如果没有重新配置下次启动时加载的配置文件，则下次启动采用缺省配置文件（如vrpcfg.zip）。如果缺省存储器中没有配置文件，则交换机启动时将使用缺省参数（即出厂配置）初始化。配置文件的文件名必须是“.cfg”或“.zip”，也必须存放在存储器的根目录下 。

补丁文件的扩展名为“.pat”，在指定下次启动时加载的补丁文件前也需要提前将补丁文件保存至交换机存储器的根目录下 。如果交换机是双主控环境，需要确保补丁文件分别保存至主用主控板和备用主控板。配置系统启动文件所用的命令如表2-13所示。

表2-13 配置系统启动文件的命令

命令	说明
startup system-software *system-file* 例如：<HUAWEI> **startup system- software** basicsoft.cc	指定交换机下次启动时所加载的系统软件。参数 *system-file*（系统软件文件名）格式为[*drive-name*] [*path*] [*file-name*]，长度范围为 4～64 个字符，不支持空格，不区分大小写。如果未指定 *drive-name*（存储器名），则此值为缺省的存储器名 如果交换机是双主控环境，还必须执行命令 **startup system-software** *system-file* **slave-board**，配置备用主控板下次启动时加载的系统软件。主用主控板和备用主控板需要指定相同版本的系统软件
startup saved-configuration *configuration-file* 例如：<HUAWEI>**startup saved- configuration** vrpcfg.cfg	指定交换机下次启动时所使用的配置文件。交换机上电时，缺省从存储器根目录中读取配置文件进行初始化。参数 *configuration-file*（配置文件名）的长度范围为 5～64 个字符，不支持空格，不区分大小写 可用 **undo startup saved-configuration** 命令取消配置的交换机下一次启动的配置文件（但需要在系统视图下执行此命令）
startup patch *file-name* [**slave-board**] 例如：<HUAWEI>**startup patch** patch.pat	（可选）可指定交换机下次启动时加载的补丁文件，**但 S2700 和 S3700 系列交换机不支持本命令**。命令中的参数和选项说明如下： （1）*file-name*：指定下次启动的补丁文件名，格式为[*drive-name*] [*path*] [*file-name*]，长度范围为 5～48 个字符，不区分大小写，不支持空格。如果未指定 *drive-name*（存储器名），则此值为缺省的存储器名 （2）**slave-board**：可选项，仅 S7700、S9700 系列交换机支持，指定备用主控板下次启动时使用的补丁文件。
display startup 例如：<HUAWEI> **display startup**	（可选）命令查看系统本次和下次启动相关的系统软件、配置文件以及补丁文件

（续表）

命令	说明
display saved-configuration [**last** | **time** | **configuration**] 例如：<HUAWEI> **display saved- configuration**	（可选）查看交换机本次或下次启动时所用的配置文件。命令中的选项说明如下。 （1）**last**：多选一可选项，显示上次保存的系统配置信息，即本次启动时使用的配置文件 （2）**time**：多选一可选项，显示最近的一次手工或者系统自动保存配置的时间。S2700、S3700 系列交换机不支持 （3）**configuration**：多选一可选项，显示设置的自动保存配置功能的参数信息，包括定时保存时间间隔、CPU 利用率等信息。S2700、S3700 系列交换机不支持 如果不带任何可选项，则直接查看交换机下次启动时所用的配置文件

【示例 1】配置下次启动使用的系统软件为basicsoft.cc。

<HUAWEI> startup system-softwarebasicsoft.cc

【示例 2】配置下次启动使用的配置文件为vrpcfg.cfg。

<HUAWEI> startup saved-configuration vrpcfg.cfg

Info: Succeeded in setting the configuration for booting system.

【示例 3】在系统视图下取消下次启动时指定的配置文件。

<HUAWEI>system-view

[HUAWEI] undo startup saved-configuration

【示例 4】指定下次启动的补丁文件为patch.pat。

<HUAWEI> startup patch patch.pat. . . . .

....

Info: Succeeded in setting main board resource file for system.

【示例 5】显示本次及下次启动相关的文件名。输出信息字段说明如表2-14所示。

<HUAWEI> display startup

MainBoard:

Configured startup system software: flash:/basicsoftware.cc

Startup system software: flash:/basicsoftware.cc

Next startup system software:　flash:/basicsoftware.cc

Startup saved-configuration file:　flash:/vrpcfg.zip

Next startup saved-configuration file:　flash:/vrpcfg.zip

Startup paf file:　NULL

Next startup paf file:　NULL

Startup license file:　NULL

Next startup license file:　NULL

Startup patch package:　NULL

Next startup patch package:　NULL

表2-14 display startup命令输出信息字段说明

项目	描述
Configured startup system software	指定的系统软件文件
Startup system software	本次启动所使用的系统软件文件
Next startup system software	下一次系统启动所使用的系统软件文件
Startup saved-configuration file	本次启动时使用的配置文件
Next startup saved-configuration file	下一次启动时使用的配置文件
Startup paf file	本次启动所使用的 PAF 文件。“NULL”表示交换机无 PAF 文件
Next startup paf file	下一次启动所使用的 PAF 文件。“NULL”表示交换机无 PAF 文件
Startup license file	本次启动所使用的License文件。“NULL”表示交换机无License文件
Next startup license file	下一次启动所使用的 License 文件。“NULL”表示交换机无 License 文件
Startup patch package	本次启动所使用的补丁文件。“NULL”表示没有指定补丁文件
Next startup patch package	下一次启动所使用的补丁文件。“NULL”表示没有指定下次启动时加载的补丁文件

【示例 6】查看下次启动时所加载的配置文件的内容。

```
<HUAWEI>display saved-configuration
#
sysname HUAWEI
#
vlan batch 1 to 10
#
cluster enable
#
……
```

【示例 7】查看本次启动时所加载的配置文件的内容。

```
<HUAWEI>display saved-configuration last
#
sysname HUAWEI
#
cluster enable
```

#

……

2.6.2 重新启动交换机

为了使指定的系统软件及相关文件生效，需要在配置完系统启动文件后，对交换机进行重新启动。重新启动交换机有以下两种方式。

（1）立即重新启动交换机：执行命令行后立即重新启动，通过设备电源开关重启，但是一般不建议这么做。

（2）定时重新启动交换机：可以设置在未来的某一时刻重新启动交换机。配置完下次系统启动文件后，为了不影响当前交换机的运行，可以将交换机设置在业务量少的时间点定时重新启动。

交换机每一次重新启动或某一单板复位的相关信息都会被详细记录下来，包括重新启动的次数、详细信息以及原因等，可以通过display reset-reason命令进行查看。在重新启动交换机之前，如果需要将当前配置在重新启动交换机后仍生效，请先确保当前配置已保存。保存配置文件的方法参见2.5.1节。

1．立即重新启动交换机

要立即重启交换机，只需在用户视图下执行reboot [fast | savediagnostic- information] 命令。命令中的两个选项说明如下。

（1）fast：二选一可选项，表示快速重启交换机，不会提示是否保存配置文件，未保存的配置信息将丢失。

（2）save diagnostic-information：二选一可选项，表示系统在重新启动前会将诊断信息保存到交换机存储器的根目录下。本可选项部分机型不支持，具体可以参考对应的产品手册。

如果执行不带任何可选项的reboot命令，则系统重启前将提示用户是否保存配置。

【示例 1】以不带任何选项的 reboot 命令重新启动交换机。重启前如果有未保存的配置，系统会提示是否保存。

```
<HUAWEI>reboot
Warning: The configuration has been modified, and it will be saved to the next s
tartup saved-configuration file cfcard:/204.cfg. Continue? [Y/N]:y
Info: If want to reboot with saving diagnostic information, input 'N' and then e
xecute 'reboot save diagnostic-information'.
System will reboot! Continue?[Y/N]:y
```

【示例 2】快速重新启动交换机，不提示是否保存配置，直接重启。

```
<HUAWEI> reboot fast
```

2．定时重新启动交换机

如果要设置定时重启，可在用户视图下使用 schedule reboot { at time |delay interval [force] }命令使能定时重新启动功能，并设置重启时间。命令中的参数和选项说明如下。

（1） time：二选一参数，设置交换机定时重新启动的具体时间。格式为 hh:mm YYYY/MM/DD，表示年月日，必须大于交换机的当前时间，且与当前时间的差值范围小于720小时（即30天）。

（2）interval：二选一参数，设置交换机在定时重新启动前等待的时间。格式为hhh:mm或mmm，其中hhh表示小时，取值范围是0～720，mm表示分钟，取值范围是0～59，mmm表示分钟，取值范围是0～43200

（3）force：可选项，指定定时强制重启交换机。如果不指定本可选项，系统首先会将当前配置与配置

文件进行比较，如果不一致，则会提示是否保存当前配置，用户进行选择后系统又将提示用户确认设置的定时重启时间，按下“Y”或者“y”键后，设置生效。如果指定了本可选项，则系统不会出现任何提示，设置生效后，当前配置不会被比较及保存，直接重启。

【示例 3】设置交换机在当天晚上22：00重新启动。

<HUAWEI> schedule reboot at 22:00

Info: The system is now comparing the configuration, please wait.

Warning: All the configuration will be saved to the configuration file for the n

ext startup:cfcard:/vrpcfg.zip, Continue?[Y/N]:y

Info: Reboot system at 22:00:00 2010/11/14(in 2 hours and 2 minutes)

confirm? [Y/N]: y

如果配置了定时重启功能，可以执行display schedule reboot命令查看交换机定时重启的相关配置。如下所示为查看当前交换机上定时重新启动的配置信息为 2013 年 4 月30日的0点重启。

<HUAWEI>display schedule reboot

Info:System will reboot at 00:00:00 2013/04/30 (in 23 hours and 53 minutes).

第3章　VRP系统登录及远程文件管理

通过上一章的学习，我们已对华为VRP系统的一些基础知识和使用方法有了比较全面的了解，本章将进一步介绍华为VRP系统。

华为VRP是一个多用户网络操作系统，不仅可以创建多个用户，通过“用户界面”区分不同的用户类型，而且可以通过用户级别为具体用户配置不同的服务和权限，就像我们非常熟悉的Windows系统中的用户一样。另外，华为公司又为用户访问交换机VRP系统提供了多种不同的登录方法，包括通过Console口本地登录和通过Telnet、STelnet、HTTP和HTTPS协议的远程登录，以满足于各种网络环境下用户的VRP系统登录需求。还可以通过各种文件传输协议（如FTP、SFTP、SCP、FTPS）进行远程文件管理，包括交换机VRP系统软件、配置文件等的上传与下载。

本章主要介绍VRP系统的各种用户界面、各种登录方法以及通过各种文件传输协议进行的远程文件管理配置方法。

3.1 VRP系统首次登录

因为华为交换机在出厂时只配置了一些基本的缺省配置，各用户的实际网络环境和网络配置也不尽相同，所以没有也不可能根据不同用户的需求进行特色配置。这就需要用户在购买新的交换机后首先了解一些VRP系统登录方法和基本系统配置。这里首先涉及的就是VRP系统的首次登录。

VRP系统的首次登录必须采用本地登录方式，因为此时交换机还没配备用于远程登录的用户、IP地址等必需配置。华为S系列交换机可以通过Console口或MiniUSB口（仅部分产品支持，如S5700LI、S5700S-LI、S5710EI系列）进行首次登录，实现对新交换机的基本配置。

3.1.1 通过Console口登录

华为S系列交换机都会提供一个Console口（基本上均为RJ-45接口类型，在接口下面会有一个CONSOLE字样），用户终端（如PC机）的串行端口（俗称COM口）可以通过随机提供的专门 Console 电缆与交换机的 Console 口直接连接。当然，事先要求在用于本地登录交换机VRP系统的PC机上安装好终端仿真软件，如Windows系统自带的超级终端软件，还有许多第三方终端仿真软件，如SecureCRT、Putty等。如果你当前使用的是Linux操作系统，则可使用各发行版本Linux系统中的minicom或者gtkterm程序作为超级终端软件。

下面仅以Windows XP系统自带的超级终端作为终端仿真软件来介绍通过 Console口本地登录华为S系列交换机VRP系统的具体操作步骤。

（1）使用产品随机附带的Console通信电缆的DB9（孔）插头插入PC机的9芯（针）串口，再将电缆的

另一端RJ-45插头端插入交换机的RJ-45 Console口中，如图 3-1所示。

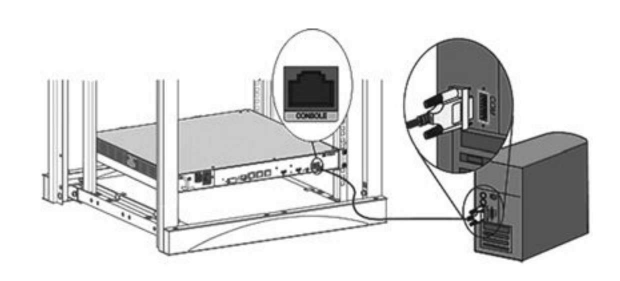


图3-1 通过Console电缆连接终端PC的COM口与交换机的Console口

（2）先后开启PC机和交换机电源，在PC机的Windows XP操作系统进入后按“开始 >所有程序 >附件 >通信 >超级终端”顺序打开超级终端软件，打开如图3-2 所示的“连接描述”对话框，在这里可以新建一个通信连接。

（3）在“名称”文本框中输入新的连 接名称（假设起名为COMM1），然后单击“确定”按钮，打开如图3-3所示的“连接到”对话框。

说明

由于目前大部分笔记本电脑没有 COM 口，只能使用 USB 接口连接。这时需 要购买一条USB-Serial电缆，其中的COM母头直接连接随设备配带的Console电缆COM 公头，USB 口连接到笔记本电脑的 USB 接口。然后在笔记本电脑中安装随USB-Serial电缆自带的驱动程序（也可在网上下载一个USB转RS-232的驱动程序）。 这样就需要在图 3-3 所示的对话框“连接时使用”下拉列表中选择由 USB 口转换生成的逻辑COM口。

（4）在“连接时使用”下拉列表中选择连接PC机的COM口（此处为COM1口），然后单击“确定”按钮打开如图 3-4 所示对话框，设置 COM 口属性。其实在这里只需单击“还原为缺省值（R）”按钮，按如下缺省设置即可：每秒位数：9600；数据位：8位；奇偶校验：无；停止位：1位；数据流控制：无。



图3-2 “连接描述”对话框



图3-3 “连接到”对话框



图3-4 “COM1属性”对话框

（5）单击“确定”按钮，超级终端就会开始与所连接的交换机建立连接，直到系统出现如下配置密码提示。

说明

首次登录时会提示用户配置登录密码（系统会自动保存此密码配置）。密码为6～16个字符，区分大小写 。为保证安全性，建议输入的密码至少包含以下几种类型 ：大写字母、小写字母、数字及特殊字符，但不能包括“？”和空格 。此处采用的是隐式方式输入，所以所输入的密码不会在终端屏幕上显示 。

如果交换机在出厂时已有初始密码，请输入初始密码“Admin@huawei.com”进入系统，但此密码不是安全密码，建议及时修改，修改方法请参见本章3.4.4节介绍。

Please configure the login password (6-16)　#---配置6～16位的登录密码

Enter Password:　#---配置密码

Confirm Password:　#---重复输入一次上述配置的密码

密码配置成功后，当用户采用密码验证方式再次通过此Console用户界面登录VRP系统时，所要输入的用户密码即为这里所配置的验证密码。有关 Console 登录的验证方式配置方法参见本章后面的3.4.4节。

配置好密码后就成功进入VRP系统了。此时用户可以键入命令对交换机进行配置，需要帮助可以随时键入“?”。

3.1.2 通过MiniUSB口登录

对于S5700LI/5700S-LI/5710EI系列交换机，如果用户PC机没有可用串口，还可以使用PC机上的USB口连接到交换机的MiniUSB口实现VRP系统的首次登录。但在通过MiniUSB口登录交换机前需要在用户终端安装MiniUSB口的驱动程序。且同一时刻，交换机上的MiniUSB口和Console口只有一个可以使用，因为它们使用的是同一条线路。

下面同样以Windows XP系统的超级终端作为终端仿真软件为例进行介绍。具体步骤如下。

（1）使用MiniUSB线缆将PC的USB口和交换机的MiniUSB口连接。

（2）双击获取交换机的MiniUSB口驱动程序，打开如图3-5所示的欢迎对话框，开始在PC端安装MiniUSB口的驱动程序。

说明

交换机 MiniUSB 口驱动程序可以从华为公司企业业务网站（ http://support. huawei.com/enterprise）中获取。登录后，在“软件下载 >产品软件 >企业网络 >交换机 >园区交换机”路径下根据产品型号和版本名称到相应路径下获取交换机的 MiniUSB口的驱动程序 Switch-MiniUSB-driver.001.zip。目前此驱动程序仅支持Windows XP/VISTA/Windows7操作系统。

（3）单击“Next”按钮，打开如图 3-6所示对话框。选择“I accept the terms of the License Agreement”单选项。



图3-5 MiniUSB口驱动程序安装欢迎对话框

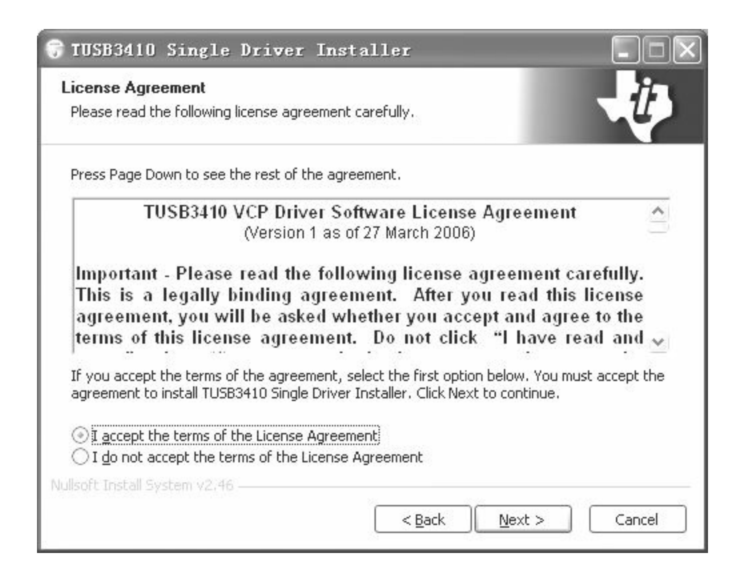


图3-6 接受软件协议对话框

（4）单击“Next”按钮，打开如图3-7所示对话框。

（5）单击“Browse”按钮可以更改驱动程序解压的路径，然后单击“Install”按钮， 对驱动程序进行解压，完成后打开如图3-8所示对话框。单击“Finish”按钮完成程序解压。

（6）在第 4步中指定的解压路径下找到“TUSB3410 Single Driver Installer”文件夹，找到并双击“setup.exe”图标，然后按提示一步步完成驱动程序的安装。安装完成后会在“交换机管理器”的“端口（COM和LPT）”栏中显示“TUSB3410 Device”，表示为已正确安装。

（7）使用终端仿真软件通过 MiniUSB 口登录交换机，这后面的登录步骤与上节介绍的通过Console口登录交换机的步骤就完全一样了，参见即可。只不过这里在如图3-3所示的对话框中所要选择的COM口不再是PC终端上的物理COM口，而是由TUSB3410口根据当前连接所创建的逻辑COM口，通常会在对话框中自动显示，不需要再更改。

最后也是可以配置验证密码，完成后进行 VRP系统，可以开始 VRP系统的使用了。

图3-7 选择驱动程序解压路径对话框



图3-8 解压完成对话框

3.2 交换机基本配置的配置

在配置业务之前，用户往往需要根据系统运行时的环境要求进行一些基本配置，以满足运行维护需求。通过以上介绍的本地登录就可以配置这些交换机的基本配置了，如交换机的时间和日期、交换机的名称和IP地址、标题文本、命令级别和用户级别切换密码等。因为命令级别和用户级别切换密码已在上一章2.2.3节详细介绍，故不再赘述，在此仅介绍前面三部分基本配置。

3.2.1 配置交换机时间和日期

华为S系列交换机与任何网络交换机一样，时间和日期的显示是非常重要的，否则交换机无法进行一些同步操作，也无法正确显示一些日志记录信息等，甚至无法正常工作。华为S系列园区交换机的时间和日期设置命令如表3-1所示（都是在用户视图下进行的 ）。

表3-1 S系列交换机的时间和日期配置命令

命令	说明
clock timezone *time-zone-name* { **add** | **minus** } *offset*	设置所在时区。命令中的参数和选项说明如下。 （1）*time-zone-name*：指定时区名称，长度范围为 1～32 个字符，区分大小写，不支持空格。我国的时区名通常写成 BJ，国际标准时区为 UTC（Universal Time Coordinated，世界协调时间） （2）**add**：二选一选项，指定与通用协调时间 UTC 相比，*time-zone-name* 时区增加的时间偏移量。即在系统缺省的 UTC 时区的基础上，加上 *offset* 参数值就可以得到 *time-zone-name* 时区所标识的时区时间。我国的时区相对 UTC 时区来说，必须加上 8 个小时的时间偏移量 （3）**minus**：二选一选项，指定将在 UTC 标准时间的基础上减去指定的时区偏移量。即在系统缺省的 UTC 时区的基础上，减去 *offset* 参数值就可以得到 *time-zone-name* 所标识的时区时间 （4）*offset*：指定与 UTC 的时间差。格式是 *HH:MM:SS*。*HH* 表示小时：如果本地时间快于 UTC 时间，取值范围为 0～14 的整数，如果本地时间慢于 UTC 时间，取值范围为 0～12 的整数；*MM* 和 *SS* 分别表示分和秒，取值范围为 0～59，当 *HH* 取值为最大值的时候，*MM* 和 *SS* 只能取值为 0 如果没有指定时区名称，则系统采用缺省值“DefaultZoneName”，可用 **undo clock timezone** 命令将本地时区恢复为缺省的 UTC 时区（通常为伦敦时区）
clock datetime *HH:MM:SS YYYY-MM-DD*	设置当前时间和日期。命令中的参数说明如下。 （1）*HH:MM:SS*：指定交换机当前时钟。*HH* 表示小时，取值范围为 0～23 的整数；*MM* 表示分钟，取值范围为 0～59 的整数；*SS* 表示秒，取值范围为 0～59 的整数 （2）*YYYY-MM-DD*：指定交换机当前年、月、日。*YYYY* 表示年份，取值范围为 2000～2099 的整数；*MM* 表示月份，取值范围为 1～12 的整数；*DD* 表示日期，取值范围为 1～31 的整数 【注意】当时区为 0 时，通过本命令设置的时间将被认为是 UTC 时间。建议设置当前时间时务必清楚所在时区，设置正确的 UTC 时间，以保证本地时间正确
clock daylight-saving-time *time-zone-name* **one-year** *start-time start-date end-time end-date offset* 或 **clock daylight-saving-time** *time-zone-name* **repeating** *start-time* { { **first** | **second** | **third** |**fourth** | **last** } *weekday month* | *start-date1* } *end-time* { { **first** | **second** | **third** | **fourth** | **last** }*weekday month* | *end-date1* } *offset* [*start-year* [*end-year*]]	（可选）设置系统夏令时，现在我们国家好象基本不用夏令时了，所以这项配置仅作一般了解即可。缺省情况下，系统没有设置夏令时。两个命令中的参数说明如表 3-2 所示。 缺省情况下，系统未使能夏令时，可用 **undo clock daylight-saving-time** 命令取消夏令时设置 当当前时间处在夏令时时，执行命令 **clock timezone** *time-zone-name* { **add** | **minus** }*offset* 设置时区名是可以成功的。但此时执行命令 **display clock** 显示的时区名为夏令时名，当夏令时结束之后，就会显示之前设置的时区名

表3-2 clock daylight-saving-time命令参数和选项说明

参数或选项	参数或选项说明
time-zone-name	指定夏令时区名称，长度范围为 1～32 个字符
one-year	指定采用一年制绝对夏令时
repeating	指定采用周期制夏令时

（续表）

参数或选项	参数或选项说明
start-time	指定起始时间，格式为 *HH:MM*，24 小时制，其中 *HH* 表示小时，取值范围为 0～23 的整数，*MM* 表示分钟，取值范围为 0～59 的整数。可以不输入 *MM*，表示 0 分，但至少需要输入一位数的 *HH* 的值，例如输入 0，则表示 0 小时 0 分
start-date	指定起始日期，格式为 *YYYY-MM-DD*，*YYYY* 表示年，取值范围为 2000～2099 的整数，*MM* 表示月，取值范围为 1～12 的整数，*DD* 表示日期，取值范围是 1～31 的整数
end-time	指定结束时间，格式为 *HH:MM*，24 小时制，其他与前面的 *start-time* 参数说明一样
end-date	指定结束日期，格式是 *YYYY-MM-DD*，其他与前面的 *start-date* 参数说明一样
first	多选一选项，指定夏令时起始或结束时间中的第一个工作日（由后面的 *weekday* 参数指定）
second	多选一选项，指定夏令时起始或结束时间中的第二个工作日（由后面的 *weekday* 参数指定）
third	多选一选项，指定夏令时起始或结束时间中的第三个工作日（由后面的 *weekday* 参数指定）
fourth	多选一选项，指定夏令时起始或结束时间中的第四个工作日（由后面的 *weekday* 参数指定）
last	多选一选项，指定夏令时起始或结束时间中的最后一个工作日（由后面的 *weekday* 参数指定）
weekday	指定夏令时起始或结束时间中的工作日，取值：Mon、Tue、Wed、Thu、Fri、Sat、Sun，分别表示从星期一到星期日
month	指定夏令时起始或结束时间中月份，取值：Jan、Feb、Mar、Apr、May、Jun、Jul、Aug、Sep、Oct、Nov、Dec，分别表示从 1 月份到 12 月份
start-date1	二选一参数，指定夏令时开始日期，格式是 *MM-DD*，*MM* 表示月，取值范围为 1～12 的整数，*DD* 表示日期，取值范围为 1～31 的整数
end-date1	二选一参数，指定夏令时结束日期，格式是 *MM-DD*，其他与前面的 *start-date1* 参数说明一样
offset	指定采用夏令时的时差（或偏移值），格式是 *HH:MM*，24 小时制。*HH* 表示小时，取值范围为 0～23 的整数，*MM* 表示分，取值范围为 0～59 的整数。可以不输入 *MM*，表示 0 分，但至少需要输入一位数的 *HH* 的值
start-year	可选参数，指定开始年份，格式是 *YYYY*，*YYYY* 取值范围为 2000～2099 的整数。如果不指定本可选参数，则表示为当前年份
end-year	可选参数，指定结束年份，格式是 *YYYY*，*YYYY* 取值范围为 2000～2099 的整数。如果不指定本可选参数，则表示为当前年份

【示例 1】假设地理位置在中国北京，设置本地时区名称为BJ，时差增加8。

<HUAWEI> clock timezone BJ add 08:00:00

【示例 2】设置系统当前日期为2013年5月1日0时0分0秒。

<HUAWEI> clock datetime 0:0:0 2013-05-01

【示例 3】按周期设置夏令时。从2013年1月的第一个星期天0点开始到2013年的4月的第一个星期天的时差为2个小时。

<HUAWEI> clock daylight-saving-timebj repeating 0 first sun jan 0 first sun apr 2 2013 2013

【示例 4】按日期设置周期夏令时。从当年1月1日的12时11分到3月4日的1时的时差值为1个小时。

<HUAWEI> clock daylight-saving-time bj repeating 12:11 1-1 1:0 3-4 1

【示例 5】设置绝对夏令时。从2013年10月2日的12时11分到2013年11月4日的1时的时差为1个小时。

<HUAWEI> clock daylight-saving-time bj one-year12:11 2013-10-2 1:00 2013-11-4 1

3.2.2 配置交换机名称和IP地址

这里所配置的IP地址可以看成是管理IP地址，专为日后进行Telnet之类的远程交换机登录使用。但这里要注意的是，在华为交换机中除了一些交换机提供的专门管理口（通常为Ethernet0/0/0接口）外，不能直接在物理接口上配置**IP**地址，仅可在VLAN接口、Loopback、Tunnel、子接口等这些逻辑接口上配置 IP 地址。管理 IP 地址通常是在管理接口，或者 VLAN 接口上配置，如果交换机没有提供管理接口，通常是在VLAN接口上配置，可直接在缺省的VLAN1接口上配置。具体配置步骤如表3-3所示。

表3-3 交换机名称和IP地址的配置步骤

步骤	命令	说明
1	**system-view** 例如：<HUAWEI> **system-view**	进入系统视图
2	**sysname** *host-name* 例如：[HUAWEI] **sysname** SWA	设置交换机名称，为 1～246 个字符，**支持空格**，**区分大小写**。当网管工具需要获取交换机的网元名称，可通过 **sys-netid** *netid* 命令设置交换机的网元名称，参数 *netid* 为长度范围为 16～240 个字符，**支持空格**，**区分大小写** 缺省情况下，华为的交换机缺省主机名为“HUAWEI”或“Huawei”，在此以“HUAWEI”为例
3	**interface** *interface-type interfa-ce-number* [SWA] **interface** vlanif 2	键入要配置 IP 地址的接口，进入接口视图。除了交换机的管理口，也可以在交换机的其他三层接口（如 VLANIF 接口）配置 IP 地址，进行 Telnet 登录
4	**ip address** *ip-address* { *mask* | *mask-length* } [**sub**] 例如：[SWA-Vlanif2] **ip address** 10.1.1.2 8	为以上接口配置 IP 地址。命令中的参数和选项说明如下。 （1）*ip-address*：指定接口的 IP 地址 （2）*mask*：二选一参数，指定所设置的 IP 地址对应的子网掩码 （3）*mask-length*：二选一参数，指定所设置的 IP 地址对应的子网掩码长度 （4）**sub**：可选项，指定以上的 IP 地址为从 IP 地址，如果不选择此可选项，则设置的 IP 地址为主 IP 地址 如果要利用这个 IP 地址进行 Telnet 远程登录的话，则还需要确保登录终端与交换机间的路由可达

【示例 】配置VLANIF2接口的主IP地址为10.1.1.2，从IP地址为11.1.1.3，子网掩码均为255.0.0.0。

<HUAWEI>system-view

[HUAWEI] interfacevlanif 2

[HUAWEI-Vlanif2] ip address 10.1.1.2 8

[HUAWEI-Vlanif2] ip address 11.1.1.3 255.0.0.0 sub

3.2.3 设置标题文本

如果需要对登录的用户提供警示或说明信息，可以设置登录交换机时或登录成功后的标题文本，但这通常是不需要进行的 。标题文本是用户在连接到交换机、进行登录验证以及开始交互配置时系统显示的一段提示信息。

标题文本有两种，一种是在登录时显示的文本，另一种是在登录成功后显示的文本。它们的配置方法都很简单，只需在系统视图下使用 **header login** { **information** text | **file** file-name }命令设置用户登录时显示的标题文本（但需要用户设置通过验证方式登录到交换机上，否则系统不会显示），使用**header shell** { **information** text | **file** file-name }命令设置用户登录成功后显示的标题文本即可。两命令中的参数说明如下。

（1）text：二选一参数，指定标题信息和内容，最多为480个字符，支持空格和换行 。标题信息以第一个英文字符作为起始符号（通常是&或者%之类的非字母类符号），最后一个英文字符作为结束符，起始符与结束符必须相同， 否则系统将会提示错误。

说明

标题信息和内容可以采用“交互输入”和“非交互输入”两种方式。交互输入就 是只需要执行不带任何参数的**header login information** 或**header shell informatio** n命令时，在自动提示下输入标题信息，用户可以根据标题信息的内容随时按回车键进行换行输入；非交互输入是指输入带text参数的以上命令。

（2）file-name：二选一参数，指定标题信息所使用的文本文件名，是通过调用文件来实现的，文件名为1～256个字符。文件名必须包含标题文件所在的绝对路径，格式为[drive] [path] [file name] ，drive是指存储器名称，path是指文件存放的目录路径，file name为文件名，且该文件名所标识的标题文件内容最大为2kB，否则配置失败。

说明

如果是在用户成功登录交换机后修改指定文件中的标题信息，系统中已经显示的 标题信息不会修改，且即使当前用户退出交换机后再重新登录交换机，标题信息也不会改变 。仅当发生以下两种情况：①交

换机重启前重新执行本命令，用户退出交换机后再重新登录交换机，标题信息将更新为修改后的标题信息；②交换机重启后，标题信息将更新为修改后的标题信息。

缺省情况下，用户登录时和登录成功时在终端上均不显示标题信息，分别可用 **undo header login** 或**undo header shell** 删除用户登录时或登录成功时在终端上显示的标题信息。

以上两命令可多次执行，如果多次配置标题信息，以最后一次配置为准（通过这种方式可以非常简单地修改标题信息，不用执行对应的**undo** 命令）。设置登录标题后，所有用户登录到系统上都能看到该标题信息。

【示例 1】配置会话建立标题（非交互方式）。

<HUAWEI>system-view

[HUAWEI] header shell information&Hello! Welcome to system!&　#---在起始字符“&”后直接输入标题信息，再以“&”作为结束字符，按回车键执行完毕

此时用户登录成功后会显示以上配置的Shell标题。

Hello! Welcome to system!

【示例 2】配置会话建立标题（交互方式）。

<HUAWEI>system-view

[HUAWEI] header shell information %　#输入起始字符“%”后按回车键，进入交互过程（在下面首先会自动显示系统提示信息，然后直接输入标题信息）

The banner text supports 480 characters max, including the start and the end cha

racter.If you want to enter more than this, use banner file instead.Input banner

text, and quit with the character '%':

Hello!

Welcome to system!%　#输入结束字符“%”后按回车键，退出交互过程，执行完毕

[HUAWEI] quit

<HUAWEI>

此时用户登录成功后会显示以上配置的Shell标题。

Hello!

Welcome to system!

【示例 3】指定登录标题使用的文件。

<HUAWEI>system-view

[HUAWEI] header login file cfcard:/header-file.txt

3.3 用户界面

当用户通过Console口、Telnet或SSH方式登录交换机时，VRP系统会为登录用户分配相应的用户界面（User-interface）管理当前用户与交换机之间的会话。华为S系列交换机的VRP系统支持“Console用户界面”和“VTY用户界面”这两大类。当用户通过Console口或者 MiniUSB 口登录交换机实现本地维护时，可以根据使用需求或对交换机安全的考虑，配置相应的Console用户界面属性；当用户通过Telnet或SSH方式登录交换机实现本地或远程维护时，可以根据用户使用需求以及对交换机安全的考虑，配置VTY用户界面。

3.3.1 用户界面概述

用户界面视图是VRP系统提供的一种命令行视图，用来配置和管理所有工作在异步交互方式下的物理接口（包括Console口和MiniUSB口）和逻辑接口（VTY虚拟接口），从而达到统一管理各种用户界面的目的。

1．Console用户界面

Console用户界面是指用户通过Console口（包括MiniUSB口）登录到交换机后的用户界面。Console 口是一种通信串行接口，由交换机的主控板提供。一块主控板提供一个 Console 口，接口类型为 EIA/TIA-232 DCE。用户终端的串行接口可以与交换机Console口直接连接，实现对交换机的本地访问。

2．VTY用户界面

VTY（Virtual Type Terminal，虚拟类型终端）是一种虚拟线路端口。用户通过终端与交换机建立Telnet或SSH连接后，即建立了一条VTY连接（或称VTY虚拟线路）。不同S系列交换机所支持的VTY连接数不一样，具体将在后面3.5节VTY用户界面配置与管理时介绍。

3．用户与用户界面的关系

用户界面与用户并没有固定的对应关系，Console类型的用户界面只有一个，但VTY类型的用户界面有多个，每个用户界面可以分配给一个用户使用。用户界面的管理和监控对象是使用某种方式登录的用户，虽然单个用户界面某一时刻可能只有一个用户使用，但它并不针对某个固定用户，因为即使是同一用户界面不同时间也可以由不同用户使用。

用户登录时，系统会根据用户的登录方式自动给用户分配一个当前空闲的、编号最小的某类型的用户界面，整个登录过程将受该用户界面视图下的配置约束。比如用户A 使用Console口登录交换机时，将受到Console用户界面视图下的配置约束；当使用VTY 1用户界面登录交换机时，将受到VTY 1用户界面视图下的配置约束。同一用户登录的方式不同分配的用户界面也不同；同一用户登录的时间不同分配的用户界面也可能不同。具体将在下节介绍。

3.3.2 用户界面的编号

华为S系列交换机上提供了多个可用的用户界面，其中Console类型的用户界面只有一个，VTY类型的用户界面有多个，而且这么多用户界面都有一个固定编号。当用户登录交换机时系统会根据此用户的登录方式自动分配一个当前空闲，且编号最小的相应类型的用户界面给这个用户。用户界面的编号包括以下两种方式。

1．相对编号

所谓“相对编号”就是针对具体类型用户界面进行的编号方式，其格式为：用户界面类型＋编号，这也是我们配置交换机功能时通常采用的编号方式。此种编号方式只能唯一指定某种类型的用户界面中的一个或一组，而不能跨类型操作。相对编号方式遵守的规则如下。

（1）Console编号：固定为CON 0，且只有这一个编号。

（2）VTY编号：第一个为VTY 0，第二个为VTY 1，最高编号为VTY 14，共有15个。

2．绝对编号

使用绝对编号方式可以唯一地指定一个用户界面或一组用户界面。可用 **display user-interface** （不带参数）命令查看交换机当前支持的用户界面以及它们的绝对编号。

每个主控板上Console口只有一个，但VTY类型的用户界面最多可有20个（其中0～14提供给普通Telnet/SSH用户的用户接口，16～20是预留给网管用户的接口，但不同交换机所支持的线路数不一样），还可在系统视图下使用**user-interface maximum-vty** 命令人为设置最大可用的用户界面个数，其缺省值为 5，即 VTY 0～4。缺省情况下， Console和VTY用户界面在VRP系统中的绝对编号和相对编号分配如表3-4所

示。

表3-4 用户界面的绝对和相对编号说明

用户界面类型	说明	绝对编号	相对编号
Console用户界面	用来管理和监控通过Console口登录的用户	0	0
VTY 用户界面	用来管理和监控通过 Telnet或SSH方式登录的用户	34～48，50～54。其中49保留，50～54 为网管预留编号	第一个为 VTY 0，第二个为 VTY 1，依此类推。缺省存在 VTY 0～4。绝对编号 34～48 对应相对编号 VTY 0～VTY 14；绝对编号 50～54 对应相对编号 VTY 16～VTY 20；其中 VTY 15 保留，VTY 16～VTY 20 为网管预留编号。只有当 VTY 0～VTY 14 全部被占用，且用户配置了 AAA 验证的情况下才可以使用 VTY 16～VTY 20

3.3.3 用户界面的用户验证和优先级

因为VRP系统是基于用户界面的网络操作系统，所以为了安全起见，需要为不同用户界面配置相应的安全保护措施，那就是配置用户界面下的用户验证。配置用户界面的用户验证方式后，用户登录交换机时VRP系统会对用户的身份进行验证。

1．用户界面的用户验证方式

VRP系统中对用户的验证方式有两种：Password验证和AAA验证：

（1）Password验证：只需要进行密码验证，不需要进行用户名验证 ，所以只需要配置密码，不需要配置本地用户。此为缺省认证方式。

（2）AAA验证：需要同时进行用户名验证和密码验证 ，所以需要创建本地用户，并为其配置对应的密码。这种方式更安全，像Telnet这样的登录方式一般是需要采用AAA验证的，但对于像SSH用户（如STelnet登录，以及SFTP、FTPS访问）需要更加严格的验证方式，如通过SSL策略中的证书、密钥认证，具体将在本章后面介绍。

2．用户界面优先级

VRP系统支持对登录用户进行分级管理，这就是我们在第2章介绍的用户级别。与命令级别一样，用户级别也对应分为0～15共16个级别，标识越高则优先级越高，具体参见第2章2.2.3节介绍。用户所能访问命令的级别由其所使用的用户界面配置的优先级（当采用不验证或者密码验证方式时）或者为用户自身配置的用户优先级别（当采用**AAA**验证方式时）决定 ，但高级别用户可以访问比他低的所有级别命令。也就是在**AAA**验证方式下，用户级别不是由所使用的用户界面级别确定 ，而是由具体的用户账户优先级别确定，更加灵活，因为这样一来，同一用户界面下的不同用户的用户级别可能不一样。当然，这也决定了在**AAA**验证方式下，必须为具体的用户配置具体的用户优先级 。

3.4 Console用户界面配置与管理

当用户通过 Console 口登录交换机实现本地维护时，可以根据实际需求配置相应的Console用户界面属性，包括Console用户界面的物理属性、终端属性、用户优先级和用户验证方式等。但这些参数都不是必须要配置的 ，用户可以结合实际需求和安全性考虑选择配置。但在配置Console用户界面之前，先需要通过终端才可以登录交换机。

另外，以下配置内容没有严格的先后顺序，仅为方便描述和理解采用步骤方式介绍。

3.4.1 配置Console用户界面的物理属性

Console用户界面的物理属性包括Console口的传输速率、流控方式、校验位、停止位和数据位。其实这些都是针对串口通信的一些属性进行配置的。具体配置步骤如表3-5所示（一般无需配置，直接采用缺省配置即可）。但要注意的是，如果改变了缺省配置，则在超级终端软件中所设置的下列属性一定要和表3-5中所设置的物理属性保持一致，否则无法登录。

表3-5 Console用户界面物理属性的配置步骤

步骤	命令	说明
1	**system-view** 例如：<HUAWEI> **system-view**	进入系统视图
2	**user-interface console** *interface-number* 例如：[HUAWEI] **user-interface console** 0	进入Console用户界面视图，参数 *interface-number* 用来指定Console口编号，只能为0

（续表）

步骤	命令	说明
3	**speed** *speed-value* 例如：[HUAWEI-ui-console0] **speed** 38400	设置 Console 用户界面的传输速率。参数 *speed-value* 用来指定 Console 用户界面的传输速率，单位为 bit/s。取值可以为：300、600、1 200、4 800、9 600、19 200、38 400、57 600 或 115 200 缺省情况下，传输速率为 9 600bit/s，可用 **undo speed** 命令恢复 Console 用户界面的传输速率为缺省值
4	**flow-control** { **hardware** | **none** | **software** } 例如：[HUAWEI-ui-console0] **flow-control hardware**	设置 Console 用户界面的流控方式。命令中的选项说明如下。 （1）**hardware**：多选一选项，指定采用硬件流控方式（采用专门的串口通信电缆连接，通过终端的主板相应芯片控制） （2）**none**：多选一选项，指定不进行流量控制 （3）**software**：多选一选项，指定采用软件流控方式（采用数据链路层协议进行流量控制） 缺省情况下，流控方式为 **none**，可用 **undo flow-control** 命令恢复流控方式为缺省的 **none** 方式
5	**parity** { **even** | **mark** | **none** | **odd** | **space** } 例如：[HUAWEI-ui-console0] **parity space**	设置 Console 用户界面的校验位。命令中的选项说明如下。 （1）**even**：多选一选项，指定采用偶校验。采用此种校验方式时，校验位的值是通过确保每个字节中的"1"的位数为偶数计算得出的 （2）**mark**：多选一选项，指定采用 Mark 校验。采用此种校验方式时，校验位始终为 1 （3）**none**：多选一选项，指定不进行校验，即无校验位 （4）**odd**：多选一选项，指定采用奇校验。采用此种校验方式时，校验位的值是通过确保每个字节中的"1"的位数为奇数计算得出的 （5）**space**：多选一选项，指定采用 Space 校验。采用此种校验方式时，校验位始终为 0 缺省情况下，校验位为 **none**，那不进行校验，可用 **undo parity** 命令恢复用户界面的校验方式为缺省的 **none** 方式
6	**stopbits** { **1.5** | **1** | **2** } 例如：[HUAWEI-ui-console0] **stopbits 2**	设置 Console 用户界面的停止位。这里的"停止位"是用来间隔不同字符数据的，仅代表时隙长度。命令中的选项说明如下： （1）**1.5**：多选一选项，指定停止位为 1.5 位，表示停止位占用了 1.5 个时隙位。此时下一步的数据传输模式配置中只能选择 5 位 （2）**1**：多选一选项，指定停止位为 1 位，表示停止位占用了 1 个时隙位。此时下一步的数据传输模式配置中只能选择 7 位或 8 位 （3）**2**：多选一选项，指定停止位为 2 位，表示停止位占用了 2 个时隙位。此时下一步的数据传输模式配置中可选择 6 位、7 位或 8 位 缺省情况下，停止位为 1 位。可用 **undo stopbits** 命令恢复用户界面停止位为缺省的 1 位，对应数据位数可以是 6、7、8

（续表）

步骤	命令	说明
7	**databits** { **5** | **6** | **7** | **8** } 例如：[HUAWEI-ui-console0] **databits 6**	设置用于表示数据的位数，也即数据传输模式。四个多选一选项分别代表数据位为 5 位（用 5 位表示数据）、6 位（用 6 位表示数据）、7 位（用 7 位表示数据）、8 位（用 8 位表示数据）。缺省情况下，数据位数为 8 位，可用 **undo databits** 命令恢复数据数位为缺省的 8 位模式

3.4.2 配置Console用户界面的终端属性

除了可配置 Console 用户界面的物理属性外，还可配置 Console 用户界面的终端属性（也就是终端控制台窗口属性），包括用户超时断连功能、终端屏幕的显示行数或列数以及历史命令缓冲区大小。具体配置步骤如表3-6所示，但这些属性均为可选项配置，因为它们都有自己的缺省值 。

表3-6 Console用户界面终端属性的配置步骤

步骤	命令	说明
1	**system-view** 例如：<HUAWEI> **system-view**	进入系统视图
2	**user-interface console** *interface-number* 例如：[HUAWEI] **user-interface console** 0	进入 Console 用户界面视图，参数 *interface-number* 用来指定 Console 口编号，只能为 0
3	**idle-timeout** *minutes* [*seconds*] 例如：[HUAWEI-ui-console0] **idle-timeout** 5	设置用户连接的超时时间，即允许用户连接闲置的最长时间。参数 *minutes* [*seconds*]分别用来指定允许闲置连接的最长时间的分钟（取值范围为 0～35 791 的整数）和秒数（取值范围为 0～59 的整数）。在设定的时间内，如果连接始终处于空闲状态，系统将自动断开该连接 缺省情况下，用户界面的最长连接闲置时间为 10min。可用 **undo idle-timeout** 命令恢复超时时间的缺省值
4	**screen-length** *screen-length* [**temporary**] 例如：[HUAWEI-ui-console0] **screen-length** 25	设置终端屏幕每屏显示的行数。参数 *screen-length* 指定终端屏幕分屏显示的行数，取值范围为 0～512 的整数。取值为 0 时表示关闭分屏功能。如果同时选择可选项 **temporary**，则表示指定的是终端屏幕临时显示行数，下次登录后仍恢复为缺省值 【说明】当用户执行某一命令的输出行数比较多，用户可以改变终端屏幕每屏显示的行数，以便查看。但通常情况，无需调整终端屏幕每屏显示的行数，且不推荐设置关闭分屏功能 缺省情况下，终端屏幕显示的行数为 24 行，可用 **undo screen-length** 命令恢复缺省设置
5	**screen-width** *screen-width* 例如：[HUAWEI-ui-console0] **screen-width** 100	设置当前终端屏幕显示的列数（每个字符为一列），取值范围为 60～512 的整数。该命令仅对 **display interface description** 命令的输出信息生效，**且只对当前连接有效，用户退出后不保存设置** 缺省情况下，终端屏幕显示的列数为 80 列，可用 **undo screen-width** 命令恢复缺省设置
6	**history-command max-size** *size-value* 例如：[HUAWEI-ui-console0] **history-command max-size** 20	设置历史命令缓冲区大小，即保存的历史命令的条数，取值范围为 0～256。缺省情况下，用户界面历史命令缓冲区大小为 10 条历史命令。可用 **undo history-command max-size** 命令恢复历史命令缓冲区的大小为缺省值

（续表）

3.4.3 配置Console用户界面的用户优先级

可以配置 Console 用户界面中的用户优先级，以实现对通过 Console 口登录交换机的用户权限控制，增加通过Console口登录交换机的安全性。

前面说了，华为VRP系统中的用户优先级（也即用户级别）共分为16个级别，级别标识为0～15，标识越高则优先级越高。用户的优先级和命令的优先级是一一对应的，即用户只能使用等于或低于自己级别的命令。具体可参见本书第2章2.3.2节。Console用户界面用户优先级的配置步骤如表3-7所示。

表3-7 Console用户界面用户优先级的配置步骤

步骤	命令	说明
1	**system-view** 例如：<HUAWEI> **system-view**	进入系统视图
2	**user-interface console** *interface-number* 例如：[HUAWEI] **user-interface console** 0	进入 Console 用户界面视图，参数 *interface-number* 用来指定 Console 口编号，只能为 0
3	**user privilege level** *level* 例如：[HUAWEI-ui-console0] **user privilege level** 15	设置 Console 用户界面的用户优先级，取值范围为 0～15 的整数。缺省情况下，**S2700 和 S3700 系列交换机的 Console 用户界面的用户优先级为 15，而其他系列交换机的 Console 用户界面的用户优先级为 3** 【注意】本命令的用户优先级设置**仅对采用密码验证方式或者不验证方式通过此用户界面登录的用户生效**。对于采用 AAA 验证方式，如果用户界面下配置的用户优先级与用户名本身所配置的用户优先级（在 AAA 视图下配置）相冲突，则以用户名本身对应的用户优先级为准

3.4.4 配置Console用户界面的用户验证方式

Console用户界面提供AAA验证、密码验证和不验证3种用户验证方式。不验证是指用户无需通过验证即可通过 Console 用户界面登录交换机，但此种验证方式没有安全保证，建议配置 AAA 验证（要求同时进行用户名验证和密码验证）或密码验证方式来增加交换机的安全性。Console用户界面的用户验证方式的配置步骤如表3-8所示，一旦配置，将对所有通过 Console 口登录交换机的用户生效，所以一定要记住所配置的验证方式、验证用户名和密码。

【经验之谈】这里要特别注意的是，在网络交换机配置中会涉及许多用户名和密码，一定不要搞混，当然最好也不要用相同的用户名和密码。这方面，建议还是用一个本子记下来，保存在一个安全的地方，否则真的很难不搞混，也很难保证不忘记。

表3-8 Console用户界面的用户验证方式的配置步骤

步骤	命令	说明
1	**system-view** 例如：<HUAWEI> **system-view**	进入系统视图
2	**user-interface console** *interface-number* 例如：[HUAWEI] **user-interface console** 0	进入 Console 用户界面视图，参数 *interface-number* 用来指定 Console 口编号，只能为 0
3	**authentication-mode { aaa | password | none }** 例如：[HUAWEI-ui-console0] **authentication-mode aaa**	设置登录 Console 用户界面的验证方式。**必须配置验证方式，否则下次用户无法成功登录交换机**。当用户首次通过 Console 口登录交换机时终端会提示设置登录密码（**参见本章 3.1.1 节第 5 步**），登录交换机后用户可以使用此命令重新设置验证方式。命令中的选项说明如下。 • **aaa**：多选一选项，指定采用 AAA 验证方式 • **password**：多选一选项，指定采用密码验证方式 • **none**：多选一选项，指定不进行验证 【说明】当配置用户界面的验证方式为 **password** 时，还需要使用第 4 步的 **set authentication password** 命令配置用户界面的验证密码。此时通过 Console 用户界面登录到交换机的用户所能访问的命令级别由登录时所用的 Console 用户界面设置的所对应的级别决定，**参见 3.4.3 节** 当配置用户界面的验证方式为 aaa 时，系统将清除在 **3.1.1 节第 5 步**为切换到此用户界面所配置的密码，重新创建登录用户名，并为之配置好账户密码。登录到交换机的用户所能访问的命令级别由下面第 7 步配置的本地用户的优先级级别决定 缺省情况下，Console 用户界面采用不验证方式，可用 **undo authentication-mode** 命令恢复为不验证方式
4	**set authentication password [cipher** *password* **]** 例如：[HUAWEI-ui-console0] **set authentication password**	（可选）设置采用密码验证方式下的本地验证密码，输入的密码可以是明文或密文。**仅当采用密码验证方式时需要配置，但如果已通过 3.1.1 节第 5 步设置了密码，则不用再通过本命令设置了** 如果不指定 **cipher** *password* 可选参数时，将采用交互方式输入明文密码（输入的密码不会在终端屏幕上显示出来），为 6～16 个字符，**区分大小写，输入的密码至少包含以下几种类型：大写字母、小写字母、数字及特殊字符**。特殊字符不能包含“？”和空格；当指定 **cipher** *password* 可选参数时，既可以输入明文密码也可以输入密文密码，密文密码长度为 56，**但无论是以哪种方式输入，最终都将以密文形式保存在配置文件中** 缺省情况下，设备没有设置本地验证的密码，可用 **undo set authentication password** 命令取消配置的本地验证密码
5	**quit** 例如：[HUAWEI-ui-console0] **quit**	退出 Console 用户界面视图
6	**aaa** 例如：[HUAWEI] **aaa**	（可选）进入 AAA 视图。**仅当采用 aaa 验证方式时需要配置**

（续表）

步骤	命令	说明
7	**local-user** *user-name* ｛**password cipher** *password* |**privilege level** *level*｝* 例如：[HUAWEI-aaa] **local-user** winda **password cipher** huawei123 **privilege level** 5	（可选）配置用于 AAA 验证的本地用户名、密码和用户优先级。**仅当采用 aaa 验证方式时需要配置**。命令中的参数说明如下。 （1）*user-name*：用来配置本地用户的用户名，为 1～64 个字符，**不支持空格，不区分大小写**。如果用户名中带域名分隔符，则认为@前面的部分是用户名，后面部分是域名。如果没有@，则整个字符串为用户名，域为缺省域 （2）*password*：可多选项，指定本地用户登录密码，可以是明文或密文密码，明文密码的长度为 1～16 个字符，**不支持空格、单引号和问号，区分大小写**；密文密码的长度为 32 个字符。但无论是明文还是密文方式输入的密码均以密文方式保存在配置文件中 （3）*level*：可多选项，指定所配置的本地用户的用户优先级，取值范围为 0～15 缺省情况下，没有创建本地用户、密码和用户优先级，可用 **undo local-user** *user-name* 命令删除对应的本地用户账户
8	**local-user** *user-name* **service-type terminal** 例如：[HUAWEI-aaa] **local-user** winda **service–type terminal**	（可选）配置 AAA 验证方式下的本地用户的接入类型为 Console 用户。**仅当采用 aaa 验证方式时需要配置** 缺省情况下，本地用户可以使用所有的接入类型，可使用 **undo local-user service-type** 命令用来将本地用户的接入类型恢复为缺省配置

【示例 1】交互式设置用户界面Console的本地验证密码为huawei2012。

<HUAWEI>system-view

[HUAWEI] user-interface console0

[HUAWEI-ui-console0] set authentication password

Please configure the login password (6-16)

Enter Password:

Confirm Password:

[HUAWEI-ui-console0]

【示例 2】设置用户界面Console的本地验证密码为密文密码（cipher模式下输入明文密码设备将以密文方式保存和显示。注意不能直接手动输入密文，但可以复制粘贴方式输入）。

<HUAWEI>system-view

[HUAWEI] user-interface vty 0 4

[HUAWEI-ui-console0] set authentication password cipher

%$%$4VZtG%zew/=BWR.$gl=O$@7.F8#～<$p,B"bH:)7Y～.(B{F=8%$%$

[HUAWEI-ui-console0]

【示例 3】创建一个用于AAA验证的本地用户user1，域名为vipdomain，用户密码是admin@12345，并指定以密文形式显示。

<HUAWEI>system-view

[HUAWEI] aaa

[HUAWEI-aaa] local-user user1@vipdomain password cipher admin@12345

3.4.5 Console用户界面管理

在交换机管理中对用户和用户界面的管理是一项非常重要的工作。Console 用户界 面配置完成后，可执行下面的 **display** 命令查看当前交换机上已登录的用户和所使用的用户界面，以及当前交换机上已配置的本地用户信息；可执行下面的**kill** 命令断开与指定用户界面连接的用户。

（1）使用**display users [all]** 命令查看所有通过用户界面（包括通过VTY用户界面）登录过的用户信

息，包括当前未连接的用户。

（2）使用**display user-interface console** ui-number [**summary**] 命令查看指定Console用户界面信息，包括用户界面的绝对和相对编号、传输速率、是否通过 Modem 拨号连接、配置的用户级别、实际用户级别、验证方式等。

（3）使用 **display local-user** 命令查看交换机上已配置的所有本地用户列表摘要信息。本命令同样可用于查看通过VTY用户界面连接的本地用户。

（4）使用 **kill user-interface** 0或 **kill user-interface console** 0命令断开与指定的Console用户界面的连接。当发现有非法用户通过Console用户登录交换机时就可以使用该命令强行断开指定用户的连接。但此命令不可对当前用户进行操作，也就是要你要中断Console用户界面的用户连接，必须使用其他用户界面，如VTY用户界面去操作。

当然本命令也可用于下面即将介绍的VTY用户界面，断开指定VTY用户界面的用户连接。此时要使用**kill user-interface** { ui-number | vty ui-number1 }命令，二选一参数ui-number 用来指定对应 VTY 界面的绝对编号；二选一参数 ui-number1 用来指定对应VTY用户界面的相对编号。

3.5 VTY用户界面配置与管理

VTY是个虚拟用户界面，它不像Console用户界面那样通过物理连接实现的，而是通过网络连接建立虚拟通道实现的，而且可以建立多条VTY虚拟通道。当用户通过Telnet或SSH方式登录交换机实现本地或远程维护时，可以根据用户使用需求以及对交换机安全的考虑配置 VTY 用户界面。同样，以下配置内容没有严格的先后顺序，仅为方便描述和理解采用步骤方式介绍。

3.5.1 配置VTY用户界面的最大个数

VTY是一个虚拟界面，同一时间可以打开多个VTY界面，实现多个用户同时连接在交换机上（Console用户界面同一时间只能允许一个用户连接）。可以配置同时登录到交换机的VTY类型用户界面的最大个数，实现对并发登录用户数的限制。VTY用户界面最大个数是指登录交换机的Telnet用户和SSH用户（如采用STelnet方式登录的用户）的总和。

VTY 用户界面最大个数的配置方法很简单，仅需在系统视图下通过 **user-interface maximum-vty** number命令配置即可，取值范围为0～15个，缺省值为5。当配置**VTY**用户界面最大个数为**0**时，任何用户（包括网管用户）都无法通过**VTY**登录到交换机，一定要小心！

注意

如果要配置的 VTY 类型用户界面的最大个数小于当前在线用户的数量，则系统 提示配置失败；如果要配置的 VTY 类型用户界面的最大个数大于当前最多可以登录用户的数量（目前华为 S 系列交换机都最多支持 5 个用户），就必须为新增加的用户界面配置验证方式。

【示例 】配置VTY用户界面最大数目为10个。

<HUAWEI > system-view

[HUAWEI] user-interface maximum-vty10

3.5.2 配置VTY用户界面的基于ACL的登录限制

可以通过访问控制列表（ACL）实现对通过 VTY 用户界面的登录进行限制，有关ACL的详细介绍及配置方法参见本书第9章相关内容。在配置VTY用户界面的登录限制前，需要先在系统视图下执行acl命令创建

一个访问控制列表并进入ACL视图，然后执行rule命令增加相应访问控制列表的规则。

用户界面支持通过基本 ACL（2000～2999）来限制源 IP 地址（即访问用户的主机或网段IP地址），支持高级ACL（3000～3999）同时限制源IP地址和目的IP地址（要访问的主机或网段IP地址），以及源端口和目的端口等。通过ACL限制VTY用户界面登录时采用以下规则。

（1）当ACL的规则配置为**permit**（允许）时：

① 如果该ACL应用在 **inbound**（入）方向，则允许指定源 IP地址的其他交换机访问本地交换机。

② 如果该ACL应用在 **outbound**（出）方向，则允许本地交换机访问指定源 IP地址的其他交换机。

（2）当ACL的规则配置为**deny**（拒绝）时：

① 如果该ACL应用在 **inbound** 方向，则拒绝其他交换机访问本地交换机。

② 如果该ACL应用在 **outbound** 方向，则拒绝本地交换机访问其他交换机。

（3）当ACL未配置规则时：

① 如果该ACL应用在 **inbound** 方向，则允许任何其他交换机访问本地交换机。

② 如果该ACL应用在**outbound** 方向，则允许本地交换机访问任何其他交换机。

通过ACL限制VTY用户界面登录的具体配置步骤如表3-9所示。

表3-9 通过ACL限制VTY用户界面登录的配置步骤

步骤	命令	说明
1	**system-view** 例如：<HUAWEI> **system-view**	进入系统视图
2	**user-interface vty** *first-ui-number* [*last-ui-number*] 例如：[HUAWEI] **user-interface vty** 0 3	进入指定的 VTY 用户界面视图。命令中的参数说明如下： （1）*first-ui-number*：欲配置的第一个 VTY 用户界面，取值范围为 0 至 3.5.1 节配置的最大 VTY 用户界面编号 （2）*last-ui-number*：可选参数，指定配置的最后一个用户界面编号。选择此参数，将同时进入多个用户界面视图。*last-ui-number* 的取值要比 *first-ui-number* 取值大。如果 *first-ui-number* 参数的取值已为所配置的 VTY 用户界面最大值，则不能再选择本可选参数

（续表）

步骤	命令	说明
3	**acl** *acl-number* { **inbound** | **outbound** } 例如：[HUAWEI-ui-vty0-3] **acl** 2000 **inbound**	配置 VTY 用户界面的基于 ACL（**要事先配置好 ACL 列表**）的登录（包括呼入和呼出两个方向）限制。命令中的参数和选项说明如下。 （1）*acl-number* 用来指定要应用的 ACL 号，取值范围为 2000～3999 （2）**inbound**：二选一选项，表示对用户界面的呼入进行限制，即限制某个地址或地址段的用户登录到本地交换机 （3）**outbound**：二选一选项，表示对用户界面的呼出进行限制，即限制已经登录的用户登录到其他交换机 缺省情况下，不对呼入、呼出进行限制，可用 **undo acl** { **inbound** | **outbound** }命令取消当前配置

【示例 1】进入VTY 1用户界面。

```
<HUAWEI>system-view
[HUAWEI] user-interface vty 1
[HUAWEI-ui-vty1]
```

【示例 2】进入VTY 1～VTY 3多个用户界面。

```
<HUAWEI>system-view
[HUAWEI] user-interface vty 1 3
[HUAWEI-ui-vty1-3]
```

【示例 3】在VTY 0用户界面上限制Telnet用户对本地交换机的连接。

```
<HUAWEI>system-view
[HUAWEI] user-interface vty 0
[HUAWEI-ui-vty0] acl 2000 inbound
```

【示例 4】在VTY 0用户界面上取消对Telnet用户连接其他交换机的限制。

```
<HUAWEI>system-view
[HUAWEI] user-interface vty 0
[HUAWEI-ui-vty0] undo acl outbound
```

3.5.3 配置VTY用户界面的终端属性

与配置Console用户界面终端属性一样，也可以配置VTY用户界面的终端属性，包括用户超时断连功能、终端屏幕的显示行数或列数以及历史命令缓冲区的大小。具体配置步骤如表3-10所示。对比表3-6可以看出，本节介绍的VTY用户界面终端属性配置与这Console用户界面终端属性配置基本一样。

表3-10 VTY用户界面终端属性的配置步骤

步骤	命令	说明
1	**system-view** 例如：<HUAWEI> **system-view**	进入系统视图
2	**user-interface vty** *first-ui-number* [*last-ui-number*] 例如：[HUAWEI] **user-interface vty** 0 3	进入指定的 VTY 用户界面视图，其他参见表 3-9 中的第 2 步说明
3	**shell** 例如：[HUAWEI-ui-vty0-3] **shell**	（可选）开启终端服务，允许用户通过此界面输入命令，对交换机进行查询、配置等操作。但 Console 用户界面不支持该命令，因为它始终开启终端服务，不需要另外开启，也不能关闭缺省情况下，在所有的用户界面上启动终端服务，可用 **undo shell** 命令关闭终端服务，不允许用户通过此界面对交换机进行操作。但在 VTY 视图下配置 **undo shell** 命令后，**则此用户界面不提供 Telnet、STelnet 和 SFTP 接入服务，一定要谨慎**

（续表）

步骤	命令	说明
4	**idle-timeout** *minutes* [*seconds*] 例如：[HUAWEI-ui-vty0-3] **idle-timeout** 20	设置用户连接的超时时间，即允许用户连接闲置的最长时间，其他参见表 3-6 中的第 3 步说明
5	**screen-length** *screen-length* [**temporary**] 例如：[HUAWEI-ui-vty0-3] **screen-length** 30	设置终端屏幕每屏显示的行数，其他参见表 3-6 中的第 4 步说明
6	**screen-width** *screen-length* 例如：[HUAWEI-ui-vty0-3] **screen-width** 100	设置当前终端屏幕显示的列数（每个字符为一列），其他参见表 3-6 中的第 5 步说明
7	**history-command max-size** *size-value* 例如：[HUAWEI-ui-vty0-3] **history-command max-size** 20	设置历史命令缓冲区大小，即保存的历史命令的条数，其他参见表 3-6 中的第 6 步说明
8	**protocol inbound** { **all** | **ssh** | **telnet** } 例如：[HUAWEI-ui-vty0-3] **protocol inbound all**	配置 VTY 用户界面支持呼入连接协议。命令中的选项说明如下： （1）**all**：多选一选项，指定 VTY 类型用户界面支持所有的协议，包括 Telnet 和 SSH （2）**ssh**：多选一选项，指定 VTY 类型用户界面仅支持 SSH 协议 （3）**telnet**：多选一选项，指定 VTY 类型用户界面仅支持 telnet 协议 缺省情况下，用户界面支持的协议是 Telnet，可用 **undo protocol inbound** 命令恢复 VTY 类型用户界面支持缺省的 Telnet 协议

【示例 】设置在虚拟终端（VTY）0到4上终止终端服务。

<HUAWEI>system-view

[HUAWEI] user-interface vty0 4

[HUAWEI-ui-vty0-4] undo shell

Warning: ui-vty0-4 will be disabled. Continue? [Y/N]:y

3.5.4 配置VTY用户界面的用户优先级

与3.4.3节介绍的可以配置Console用户界面优先级一样，也可以配置VTY用户界面的用户优先级，实现对不同用户访问交换机权限的限制，增加交换机管理的安全性。用户的优先级分为16个级别，级别标识为0～15，标识越高则优先级越高。用户的优先级和命令的优先级是相对应的，即用户只能使用等于或低于自己级别的命令。

VTY用户界面优先级的配置方法也与表3-7中介绍的Console用户界面优先级配置方法基本一样，只是所在的用户界面视图不一样，要在第 2 步中使用 **user-interface vty** first-ui-number [last-ui-number] 命令进入对应的VTY用户界面，然后进行第 3步的用户界面优先级配置。缺省情况下，**VTY**用户界面对应的缺省命令访问级别是**0**（即最低的访问级，仅具有浏览权限 ）。如果用户界面下配置的用户优先级与用户名本身对应的用户优先级（在 AAA 视图下配置）冲突，以为对应用户配置的用户级别为准。具体命令说明参见表3-7。

3.5.5 配置VTY用户界面的用户验证方式

与3.4.4节介绍的Console用户界面验证方式一样，VTY用户界面也提供AAA验证、密码验证和不验证 3 种用户验证方式。不验证是指用户无需通过验证即可通过 VTY 用户界面登录到交换机，此种方式没有安全保证。建议配置 AAA 验证或密码验证方式来增加交换机管理的安全性。

VTY用户界面的用户验证方式的配置方法与3.4.4节介绍的Console用户界面的用户验证方式的配置方法也基本一样，只是所在的用户界面视图不一样，要在表3-8中的第 2步使用user-interfacevty first-ui-number [last-ui-number] 命令进入到对应的VTY用户界面。另外，在表 3-8第 8步中要使用 local-user user-name

service-type { telnet | ssh }命令配置VTY用户界面支持的服务类型，这里要选择Telnet或者SSH服务。其他参见3.4.4节的表3-8中的配置。

3.5.6 VTY用户界面管理

VTY用户界面配置完成后，请执行下面的**display** 任意视图命令检查配置结果。

（1）使用**display users [all]** 命令查看包括每个VTY用户界面下的用户登录信息。也就是可以查看当前有哪些用户连接在交换机上。

（2）使用**display user-interface maximum-vty** 命令查看当前配置（在 3.5.1节配置）的VTY类型用户界面的最大个数。

（3）使用 **display user-interface vty** ui-number1 [**summary**] 命令查看指定的用户界面信息。

（4）使用**display local-user** 命令查看本地交换机上配置的所有本地用户的属性信息。

（5）使用**display vty mode** 命令查看VTY模式，分“人机模式”（Human-Machine Mode，即人机交互模式）和“机机模式”（Machine-Machine Mode，即机机交互模式）两种。本命令**S2700**和**S3700**系列交换机不支持 。

3.6 用户登录配置与管理

介绍了用户界面的基本属性配置方法后，就要正式介绍用户通过这些用户界面登录华为S系列交换机的配置方法了。这里所说的“用户登录”是指登录到华为S系列交换机的VRP系统，通过终端软件实现本地或者远程的交换机配置与管理。它与本章后面介绍的用户访问是不同的，用户访问是指通过软件建立与交换机的网络连接，以便实现文件管理（不能进行交换机配置）。

3.6.1 用户登录概述

在华为S系列交换机中，当交换机作为服务器时，用户可以通过Console口、Telnet、STelnet（安全Telnet）或者Web方式登录本交换机。STelnet登录方式也称SSH（Secure Shell，安全外壳）登录方式，是基于SSH协议进行的。当交换机作为客户端时，可以从 本交换机通过 Telnet 或 STelnet 方式登录其他交换机，实现对网络上其他交换机的管理和维护。因篇幅原因，本章仅介绍交换机作为各种服务器，用户通过用户界面登录到 S系列交换机VRP系统的配置方法。

1．命令行登录方式

通过Console口、Telnet或STelnet方式登录交换机后，可以使用交换机提供的VRP系统命令行界面对交换机进行配置与管理。这三种登录方式都属命令行登录方式，需要配置相应登录方式的用户界面。这三种命令行用户登录方式的比较如表3-11所示。

表3-11 三种命令行登录方式的比较

登录方式	优点	缺点	应用场景	说明
通过Console口登录	使用专门的Console通信电缆连接，保证可以对交换机有效控制	不能远程登录维护交换机	当对交换机进行第一次配置时，可以通过Console口登录交换机进行配置；当用户无法进行远程登录交换机时，可通过Console口进行本地登录；当交换机无法启动时，可通过Console口进入BootROM进行诊断或系统升级	通过Console口进行本地登录是登录交换机最基本的方式，也是其他登录方式的基础
通过Telnet登录	便于对交换机进行远程管理和维护，不需要为每一台交换机都连接一个终端，极大地方便了用户的操作	传输过程采用TCP协议进行明文传输，存在安全隐患	终端连接到网络上，使用Telnet方式登录交换机，进行本地或远程的配置。应用在对安全性要求不高的网络	缺省情况下，用户不能通过Telnet方式直接登录交换机。如果需要通过Telnet方式登录交换机，可以先通过Console口本地登录交换机，并完成以下配置。 （1）确保终端和登录的交换机之间路由可达（缺省情况下，交换机上没有配置IP地址） （2）配置Telnet服务器功能及参数 （3）配置Telnet用户登录的用户界面
通过Stelnet登录	STelnet协议实现在不安全网络上提供安全的远程登录，保证了数据的完整性和可靠性，保证了数据的安全传输	配置较复杂	如果网络对于安全性要求较高，可以通过STelnet方式登录交换机。STelnet基于SSH协议，提供安全的信息保障和强大验证功能，保护交换机不受IP欺骗和明文密码截取等攻击	缺省情况下，用户不能通过STelnet方式直接登录交换机。如果需要通过STelnet方式登录交换机，可以先通过Console口本地登录或Telnet远程登录交换机，并完成以下配置。 （1）确保终端和登录的交换机之间路由可达（缺省情况下，交换机上没有配置IP地址） （2）配置STelnet服务器功能及参数 (3)配置SSH用户登录的用户界面 （4）配置SSH用户

2．Web网管登录方式

通过HTTP或HTTPS方式登录交换机时，因为交换机内置了一个Web服务器，所 以用户可以从终端（如PC机）通过Web浏览器登录到交换机，使用交换机提供的图形界面直观地管理和维护交换机。这两种登录方式都属Web网管登录方式，但必须要确保交换机上已经加载了对应版本的Web网页文件。

注意

Web网管方式虽然是通过图形界面直观地管理交换机，便于用户操作，但目前所 能提供的仅是对交换机日常维护及管理的基本功能，如果需要对交换机进行较复杂或精细的管理，仍然需要使用命令行方式。

HTTPS登录方式是将HTTP和SSL结合，通过SSL对服务器身份进行验证，对传输的数据进行加密，从而实现了对交换机的安全管理。而HTTP协议本身不能对Web服务器的身份进行验证，所以当通过HTTP方式登录交换机时存在很大的安全隐患。为了解决这一问题，在华为VRP系统通过HTTP方式登录的过程中，传输的用户名和密码也必须使用HTTPS安全协议。两种Web网管登录方式的比较如表3-12所示。

表3-12 两种Web网管用户登录方式比较

登录方式	相同点	不同点	说明
HTTP方式	都需要加载SSL证书，用于登录交换机时的身份验证。配置几乎相同（差别见不同点）	登录地址为 **http://*IP***，当用户请求登录页面时，会重定向到HTTPS登录页面（**https://*IP***），用户登录成功后，再跳转到HTTP页面，后续的数据交互仍使用HTTP协议。**在使能HTTP服务前，必须要先使能HTTPS服务**	如果需要通过HTTP或HTTPS方式登录交换机，需要完成以下配置。 （1）确保已加载了Web网页文件。 （2）配置HTTP/HTTPS服务和HTTP用户（缺省情况下，HTTP及HTTPS服务功能未使能，交换机提供缺省的HTTP用户的用户名为admin，密码为admin） 【说明】缺省情况下，Web网页文件中已经包含了SSL证书，当网页文件被加载后，用户无需进行相应的SSL策略的配置（交换机有缺省的SSL策略）。当然，也可以从CA（Certificate Authority）处重新获取数字证书，然后进行手动配置SSL策略
HTTPS方式		登录地址为 **https://*IP***，登录成功后，通过SSL对数据进行加密，安全性更高。**仅需使能HTTPS服务**	

说明

因本地的 Console 登录方式已在本章 3.1 节有详细介绍，故不再赘述，下面仅介 绍Telnet、STelnet或者Web这几种远程登录方式的配置方法。

3.6.2 配置用户通过Telnet登录交换机

Telnet协议在TCP/IP协议族中属于应用层协议，通过网络提供远程登录和虚拟终端功能。以服务器/客户端（Server/Client）模式工作，Telnet客户端向Telnet服务器发起请求，Telnet服务器提供Telnet服务。在通过Telnet登录交换机的过程中，交换机是作为Telnet服务器，当然交换机也可作为Telnet客户端向其他Telnet服务器（如其他交换机、路由器等交换机）发起连接。

1．配置任务

在配置用户通过Telnet登录交换机之前，需确保终端与交换机之间路由可达。其他Telnet登录配置任务如下。

（1）配置Telnet服务器功能和参数：包括使能Telnet服务器功能和Telnet服务器参数配置。

（2）配置Telnet用户登录的VTY用户界面：指定可用于Telnet登录的VTY用户界面，并配置相关VTY用户界面属性，包括VTY用户界面的用户优先级、验证方式、呼入限制、呼出限制等。

本项配置任务的配置方法参见本章 3.5.1～3.5.4 节（在表 3-10 中第 8 步要通过**protocol inbound telnet**命令配置对应的VTY用户界面支持Telnet服务），验证方式的配置参见本章3.5.5节，建议采用AAA验证方式（也可以采用其他验证方式 ）。

（3）配置Telnet类型的本地用户AAA验证方式：包括配置验证方式为AAA时的用户名和密码，并支持Telnet服务，具体参见本章3.5.5节介绍。如果采用的是密码验证或者不验证这两种验证方式，则不需要进行本项配置任务 。

（4）从终端通过Telnet登录交换机：从终端通过Telnet客户端软件登录到交换机VRP系统。

华为S系列交换机有关Telnet登录的参数缺省配置如表3-13所示。

表3-13 Telnet登录的相关参数缺省值

参数	缺省值
Telnet 服务器功能	使能
Telnet 服务器端口号	23
VTY 用户界面的验证方式	无验证方式
VTY 用户界面所支持的协议	Telnet 协议
用户级别	VTY 用户界面对应的缺省命令访问级别是 0

因为以上第2项和第3项配置任务在本章3.5节已有详细介绍，故下面仅介绍以上配置任务中的第1项和第

4项。

2．配置Telnet服务器功能和参数

用户终端建立与交换机的Telnet连接之前，需要首先确保交换机的Telnet服务功能已经使能。配置Telnet服务器功能的具体步骤如表3-14所示。

表3-14 Telnet服务器功能及参数的配置步骤

步骤	命令	说明
1	**system-view** 例如：<HUAWEI> **system-view**	进入系统视图
2	**telnet server enable** 例如：[HUAWEI] **telnet server enable**	（可选）使能 Telnet 服务器功能。不过，缺省情况下，Telnet 服务器功能已处于使能状态，所以也可以不进行配置步骤，除非原来人工关闭了该服务器功能。可用 **undo telnet server enable** 命令关闭 Telnet 服务器，禁止 Telnet 用户登录
3	**telnet server port** *port-number* 例如：[HUAWEI] **telnet server port** 1028	（可选）配置 Telnet 服务器的监听端口号，取值范围为 23 或 1 025～55 535 的整数。缺省情况下，监听端口号是 TCP 23。不过，重新配置 Telnet 服务器的监听端口号可使攻击者无法获知更改后的 Telnet 监听端口号，有效防止了攻击者对 Telnet 服务标准端口的登录。可通过 **undo telnet server port** 命令恢复 Telnet 服务器的监听端口号为缺省值的 TCP 23 号端口

随后按照本章3.5节的内容配置好Telnet登录中所需的VTY界面属性和Telnet登录AAA验证配置。然后就可以按照下面介绍的登录方法直接从终端Telnet登录到交换机VRP系统了。

3．从终端通过Telnet登录到交换机

从终端通过Telnet登录到交换机VRP系统，可以选择使用Windows命令行提示符或第三方软件，此处以Windows命令行提示符为例。具体步骤如下。

（1）进入Windows的命令行提示符。

（2）执行Windows命令 telnet ip-address port，通过Telnet方式登录到交换机VRP系统，如下所示。如果使用交换机上配置的Telnet服务器端口为缺省的TCP 23号端口，则不用键入端口号。

C:\Documents and Settings\Administrator> telnet 10.137.217.177 1025

按下回车键后，在提示信息下输入 AAA 验证方式配置的登录用户名和密码。验证通过后，出现用户视图的命令行提示符，至此用户成功登录交换机。

说明

根据你为Telnet登录配置的验证方式的不同，所显示的验证信息也会不同，如果 配置的是不验证方式，则不会出现验证信息，直接登录进交换机VRP系统。

```
Login authentication
Username:huawei
Password:
Info: The max number of VTY users is 8, and the number
      of current VTY users on line is 2.
      The current login time is 2012-08-06 18:33:18.
<Telnet Server>
```

4．Telnet登录管理

Telnet登录成功后，可以通过display users [all] 任意视图命令查看用户界面连接情况（选择可选项 all 时则显示所有通过用户界面登录的用户的信息，包括未连接的用户界面，否则仅显示当前已连接的用户界面下的用户信息；通过display tcp status任意视图命令查看当前建立的所有TCP连接情况；通过display telnet server status任意视图命令查看Telnet服务器的状态和配置信息。

3.6.3 通过Telnet登录交换机的配置示例

本示例基本网络结构如图 3-9 所示，PC 与交换机之间的路由可达。现要求在担当Telnet服务器的S系列交换机端配置Telnet用户，以AAA验证方式登录到VRP系统，并配置安全策略，保证只有当前管理员使用的PC（IP地址为10.1.1.1/32）才能通过指定的VTY用户界面登录交换机。

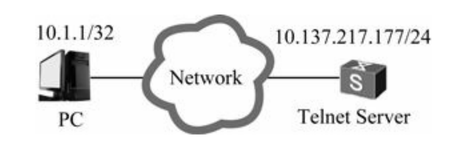


图3-9 通过Telnet登录交换机的配置示例拓扑结构

1．基本配置思路

Telnet登录方式采用的是VRP系统的VTY虚拟线路，再加上本示例要求采用AAA验证方式，并需要通过ACL控制允许通过Telnet的终端用户，所以本示例的基本配置思路如下。

（1）配置所用的VTY用户界面属性，指定支持Telnet服务的VTY用户界面，并在指定的VTY用户界面下配置ACL策略控制允许Telnet登录交换机的终端，以保证只有当前管理员使用的PC才能登录交换机的VRP系统。

（2）配置Telnet登录的AAA验证方式，并创建用于AAA验证的本地用户名和密码，以及支持的Telnet服务和命令访问级别。

（3）使能Telnet服务器，并配置Telnet服务器功能属性。

2．具体配置步骤

按照以上配置思路，即可得出如下的具体配置步骤。

（1）配置Telent登录所用的VTY用户界面的终端属性，指定VTY 0～7这8条VTY虚拟通道可用于Telnet登录。

```
<HUAWEI>system-view
[HUAWEI] sysnameTelnet Server
[Telnet Server] user-interfacevty 0 7
[Telnet Server-ui-vty0-7] shell
[Telnet Server-ui-vty0-7] idle-timeout 20
[Telnet Server-ui-vty0-7] screen-length 30
[Telnet Server-ui-vty0-7] history-command max-size20
```

（2）配置Telent登录VTY用户界面的AAA验证方式以及用户级别。

```
[Telnet Server-ui-vty0-7] authentication-mode aaa
[Telnet Server-ui-vty0-7] user privilege level 15
[Telnet Server-ui-vty0-7] quit
```

（3）配置控制通过Telnet访问交换机的用户ACL策略。

```
[Telnet Server] user-interface maximum-vty 8    #---配置VTY用户界面的最大个数
[Telnet Server] acl 2001
[Telnet Server-acl-basic-2001] rule permit source10.1.1.1 0    #---配置仅允许 IP地址为10.1.1.1的主机访问
[Telnet Server-acl-basic-2001] quit
[Telnet Server] user-interface vty 0 7
```

[Telnet Server-ui-vty0-7] acl 2001 inbound   #---在VTY 0～7这8个用户界面中应用上面的ACL

（4）创建用于Telnet登录AAA验证的用户名和密码，对Telnet服务的支持，以及用户访问级别（此示例中用户huawei最终生效的级别为3，并没有按照Telnet用户界面中所设置的15级别生效）。

[Telnet Server] aaa

[Telnet Server-aaa] local-userhuawei password cipherhello@123

[Telnet Server-aaa] local-userhuawei service-type telnet

[Telnet Server-aaa] local-userhuawei privilege level 3

[Telnet Server-aaa] quit

（5）使能Telnet服务器功能，并配置Telnet服务器的监听端口号。

[Telnet Server] telnet server enable

[Telnet Server] telnet server port1025

（6）客户端登录到的交换机VRP系统。

进入管理员PC的Windows的命令行提示符，执行相关命令，通过Telnet方式登录交换机。

C:\Documents and Settings\Administrator> telnet 10.137.217.177 1025

按下回车键后，在验证信息中按提示输入AAA验证方式配置的登录用户名和密码，验证通过后，出现用户视图的命令行提示符，至此用户成功登录交换机。

Login authentication

Username:huawei

Password:

Info: The max number of VTY users is 8, and the number

of current VTY users on line is 2.

The current login time is 2012-08-06 18:33:18.

<Telnet Server>

3.6.4 配置用户通过**STelnet**登录交换机

前面介绍的Telnet服务在传输过程采用的是TCP协议进行明文传输和缺少安全的认证方式，容易招致DoS（Denial of Service，拒绝服务）、主机 IP地址欺骗和路由欺骗等恶意攻击。本节介绍的STelnet（Secure Telnet，安全Telnet）是基于SSH（Secure Shell，安全外壳）协议，在传输过程中客户端和服务器端之间需经过协商，建立安全传输连接，以确保在登录和远程交换机配置与管理中的数据传输安全。SSH通过以下措施实现在不安全的网络上提供安全的远程登录。

（1）支持 RSA（Revest-Shamir-Adleman，这是开发这种算法的三个人的名字）和DSA（Digital Signature Algorithm，数字签名算法）加密、认证方式，RSA用于对发送的数据进行加密和数字签名，DSA仅用于对数据进行数字签名。

（2）支持用加密算法DES（Data Encryption Standard）、3DES、AES128（Advanced Encryption Standard 128）对用户名密码以及传输的数据进行加密。

华为S系列交换机支持SSH服务器功能，可以接收多个SSH客户端的连接。同时， S系列交换机还支持SSH客户端功能，可以与支持SSH服务器功能的交换机建立SSH连接，从而实现从本地交换机通过SSH登录到远程交换机。目前，交换机作为SSH服务器端时支持SSH2和SSH1两个版本，但交换机作为**SSH**客户端时只支持**SSH2**版本 。

1．配置任务

在配置用户通过 STelnet 登录交换机之前，需要确保终端与交换机之间路由可达，且在终端上已安装SSH客户端软件。其他STelnet登录配置任务如下。

（1）配置STelnet服务器功能及参数：包括服务器本地密钥对生成、STelnet服务器功能的开启以及服务器参数的配置，如监听端口号、密钥对更新时间、SSH验证超时时间或SSH验证重试次数等。

（2）配置SSH用户登录的用户界面：包括VTY用户界面的用户优先级、用户验证方式（仅可选择AAA验证模式）、支持SSH协议及其他VTY用户界面属性。

本项配置任务的配置方法参见本章 3.5.1～3.5.4 节（在表 3-10 中第 8 步要通过protocol inbound ssh命令配置对应的VTY用户界面支持SSH服务），验证方式的配置参见本章3.5.5节，建议采用AAA验证方式（也可以采用其他验证方式 ）。

（3）配置SSH用户：包括SSH用户名和密码、验证方式和服务方式等。

（4）用户通过STelnet登录交换机：从终端通过SSH客户端软件登录到交换机的VRP 系统。

华为S系列交换机有关STelnet登录的参数缺省配置如表3-15所示。

表3-15 STelnet登录的相关参数缺省值

参数	缺省值
STelnet 服务器功能	关闭
SSH 服务器端口号	22
SSH 服务器密钥对的更新周期	0 小时，表示永不更新
SSH 连接验证超时时间	60s
SSH 连接的验证重试次数	3
SSH 服务器兼容低版本功能	使能
VTY 用户界面的验证方式	无验证方式
VTY 用户界面所支持的协议	Telnet 协议
SSH 用户的验证方式	验证方式是空，即不支持任何验证方式
SSH 用户的服务方式	服务方式是空，即不支持任何服务方式
SSH 服务器为用户分配公钥	没有为用户分配公钥
用户级别	VTY 用户界面对应的缺省命令访问级别是 0

下面仅介绍以上配置任务中的第1、第3和第4项。

2．配置STelnet服务器功能和参数

Stelnet服务器功能和参数的配置步骤如表3-16所示。主要包括创建用于传输数据加密的SSH本地密钥对，使能STelnet服务器功能，配置SSH监听端口、SSH密钥更新周期、SSH验证重试次数、SSH连接超时等属性。

表3-16 STelnet服务器功能和参数的配置步骤

步骤	命令	说明
1	**system-view** 例如：<HUAWEI> **system-view**	进入系统视图
2	**rsa local-key-pair create** 例如：[HUAWEI]**rsa local-key-pair create** 或 **dsa local-key-pair create**（S2700、S3700 系列不支持） 例如：[HUAWEI] **dsa local-key-pair create**	根据所采用的加密算法，生成本地 RSA 主机密钥对和服务器密钥对，或 DSA 主机密钥对，产生的密钥对命名方式为“交换机名称_server”和“交换机名称_host”，用于服务器向客户端发送数据时的数据加密保护。 执行命令后生成的密钥对将保存在交换机中（**但不会保存在配置文件中**），交换机重启后不会丢失。密钥对生成后，可以执行 **display rsa local-key-pair public** 或 **display dsa local-key-pair public** 命令查看本地密钥对中的公钥部分信息 输入命令后，交换机会提示用户输入主机密钥的位数，RSA 服务器密钥对的位数与 RSA 主机密钥对的位数至少相差 128 位，最小长度为 512 位，最大长度为 2 048 位，缺省为 2 048 位；DSA 主机密钥对的长度可为 512、1 024、2 048 位，缺省情况下为 512 位 缺省情况下，没有配置任何本地 RSA 或 DSA 密钥对。如果原来已有 RSA 密钥对存在，则系统会提示用户确认是否替换原有密钥
3	**stelnet server enable** 例如：[HUAWEI] **stelnet server enable**	使能 SSH 服务器端的 STelnet 服务功能。缺省情况下，STelnet 服务功能未使能，可使用 **undo stelnet server enable** 命令关闭 SSH 服务器端的 STelnet 服务。**去使能 SSH 服务器的 STelnet 服务后，所有的客户端将断开连接**

（续表）

步骤	命令	说明
4	**ssh server port** *port-number* 例如：[HUAWEI] **ssh server port** 2018	（可选）配置 SSH 服务器监听端口号，取值范围为 22 或 1 025～55 535 的整数。如果配置了新的监听端口号，SSH 服务器端先断开当前已经建立的所有 STelnet 连接，然后使用新的端口号开始监听。这样可以有效防止攻击者对 SSH 服务标准端口的访问，确保安全性 缺省情况下，SSH 服务器端监听端口号是 22，可用 **undo ssh server port** 命令恢复SSH服务器端监听端口为缺省的22号端口
5	**ssh server rekey- interval** *interval* 例如：[HUAWEI]**ssh server rekey-interval** 2	（可选）配置 SSH 服务器密钥对更新周期，取值范围为 0～24 的整数小时。配置服务器密钥对更新时间，可使当 SSH 服务器密钥对的更新周期到达时，自动更新服务器密钥对，从而可以保证安全性 缺省情况下，SSH 服务器密钥对的更新时间间隔是 0，**表示永不更新**。可用 **undo ssh server rekey-interval** 命令恢复配置的 SSH 服务器密钥对更新周期为缺省值 0，即永不更新
6	**ssh server timeout** *seconds* 例如：[HUAWEI] **ssh server timeout** 100	（可选）配置 SSH 验证超时时间，取值范围为 1～120 的整数秒。当设置的 SSH 验证超时时间到达后，如果用户还未登录成功，则终止当前连接，确保了安全性 缺省情况下，SSH 连接验证超时时间是 60s，可用 **undo ssh server timeout** 命令恢复 SSH 验证超时时间为缺省的 60s
7	**ssh server authentication-retries** *times* 例如：[HUAWEI]**ssh server authentication-retries**	（可选）配置 SSH 验证重试次数，取值范围是 1～5 的整数。配置 SSH 验证重试次数用来设置 SSH 用户请求连接的验证重试次数，防止非法用户登录 缺省情况下，SSH 连接的验证重试次数是 3，可用 **undo ssh server authentication-retries** 命令恢复 SSH 验证重试次数为缺省的 3 次
8	**ssh server compatible-ssh1x enable** 例如：[HUAWEI]**ssh server compatible-ssh1x enable**	（可选）使能 SSH 服务器兼容低版本 SSH 协议，主要应用于客户端与服务器的版本协商阶段。客户端与服务器建立 TCP 连接后，开始协议版本协商，以期与服务器达成一个可以工作的协议版本 SSH 服务器按以下规则比较客户端发来的版本，决定是否能与客户端一起工作： • 如果客户端的协议版本号低于 1.3 或高于 2.0，则版本协商失败，断开连接 • 如果客户端的协议版本为大于等于 1.3 并且小于 1.99，如果系统配置为兼容 SSH1.X 方式，则进入 SSH1.5 SERVER 模块，后续进行 SSH1.x 协议流程，否则版本协商失败，断开与客户端的连接 • 如客户端协议版本为 1.99 或 2.0，则进入 SSH2.0 SERVER 模块，后续进行 SSH2.0 协议流程。 该配置在下次登录时生效。缺省情况下，SSH2.0 协议的服务器是兼容 SSH1.X 服务器功能的

【示例 1】在交换机上创建RSA本地主机密钥对和服务器密钥对。

<HUAWEI>system-view

[HUAWEI] rsa local-key-pair create

The key name will be: HUAWEI_Host

The range of public key size is (512 ～ 2048).

NOTES: If the key modulus is greater than 512,

it will take a few minutes.

Input the bits in the modulus[default = 2048] :1024

Generating keys. .

...........++++++++

...........................++++++++

....+++++++++

...+++++++++

【示例 2】在交换机上创建DSA本地主机密钥。

<HUAWEI>system-view

[HUAWEI] dsa local-key-pair create

Info: The key name will be: HUAWEI_Host_DSA.

Info: The key modulus can be any one of the following : 512, 1024, 2048.

Info: If the key modulus is greater than 512, it may take a few minutes.

Please input the modulus [default=512] :

Info: Generating keys. .

Info: Succeeded in creating the DSA host keys.

3．配置SSH用户

S系列交换机支持RSA、DSA、password、password-rsa、password-dsa和all六种用户验证方式。其中password-rsa 验证需要同时满足 password 验证和 RSA 验证；password-dsa验证需要同时满足password验证和DSA验证；all验证是指password验证、RSA或DSA验证方式满足其中一种即可。具体的配置步骤如表3-17所示。

表3-17 SSH用户的配置步骤

步骤	命令	说明
1	**system-view** 例如：<HUAWEI> **system-view**	进入系统视图
2	**ssh user** *user-name* 例如：[HUAWEI] **ssh user** winda	创建 SSH 用户，为 1～64 个字符，**不支持空格，不区分大小写**
3	**ssh user** *user-name* **authentication-type** { **password** | **rsa** |**passwo-rd-rsa** | **dsa** | **password-dsa**| **all** } 例如：[HUAWEI]**ssh user** winda **authentication- type rsa**	配置 SSH 用户的验证方式。**对于新用户必须指定其验证方式，否则用户无法登录**。但新配置的验证方式在下次登录后才生效。命令中的参数和选项说明如下。 （1）*user-name*：指定上一步创建的 SSH 用户名 （2）**password**：多选一选项，指定 SSH 用户采用密码验证方式。但要注意，**这里不仅需要进行密码验证，还需要进行用户账户名验证**。在服务器端由 AAA 为每一个合法用户分配一个用于登录时进行身份验证的口令，即在服务器端存在“用户名＋口令”的一一对应的关系。当某个用户请求登录时，服务器需要对用户名以及其口令分别进行鉴别，其中任何一项不能验证通过均告验证失败，拒绝该用户的登录请求。该验证方式的用户的账号信息可以配置在交换机或者远程验证服务器（如 RADIUS 验证服务器）上 （3）**rsa**：多选项一选项，指定 SSH 用户采用 RSA 验证方式。如果采用此种验证方式，服务器必须检查用户是否是合法的 SSH 用户，检查公钥对于该用户是否合法，检查用户数字签名是否合法。若三者同时满足，则身份验证通过；若其中任何一项不能验证通过均告验证失败，拒绝该用户的登录请求 （4）**password-rsa**：多选项一选项，指定 SSH 用户采用密码和 RSA 两种验证方式。SSH 服务器可以要求客户端进行身份验证的过程中同时进行 Publickey（公钥）身份验证和 Password 身份验证，只有在两者同时满足的情况下，才认为客户端身份验证通过

（续表）

步骤	命令	说明
3	**ssh user** *user-name* **authentication-type { password | rsa |passwo-rd-rsa | dsa | password-dsa| all }** 例如：[HUAWEI]**ssh user** winda **authentication- type rsa**	（5）**dsa**：多选一选项，指定SSH用户采用DSA验证方式（**S2700和S3700系列交换机不支持**）。DSA和RSA验证相同，服务器必须检查用户是否是合法的SSH用户，检查公钥对于该用户是否合法，用户数字签名是否合法。若三者同时满足，则身份验证通过；若其中任何一项不能验证通过均告验证失败，拒绝该用户的登录请求。但相比RSA验证，DSA验证采用数字签名算法进行加密，具有更广泛的应用性，很多工具仅支持使用DSA进行服务器和客户端验证 （6）**password-dsa**：多选一选项，指定SSH用户采用密码和DSA两种验证方式（**S2700和S3700系列交换机不支持**）。SSH服务器可以要求客户端进行身份验证的过程中同时进行Publickey身份验证和Password身份验证，只有当两者同时满足的情况下，才认为客户端身份验证通过 （7）**all**：多选项一选项，指定SSH用户密码或RSA或DSA验证方式（**如果是S2700或S3700系列交换机，则不支持DSA验证**）。可以要求客户端在进行身份验证的过程中进行公钥验证或密码验证，只要满足其中一个验证，就认为客户端身份验证通过 缺省情况下，SSH用户的验证方式为空，即不支持任何验证方式，可用**undo ssh user** *user-name* **authentication-type**命令恢复SSH用户的验证方式到缺省情况 如果没有使用本命令为相应SSH用户配置验证方式，则可以直接使用**ssh authentication-type default password**命令为该用户配置SSH验证缺省采用密码验证。在用户数量比较多时，对用户使用缺省密码验证方式可以简化配置，此时只需再配置AAA用户即可
4	**ssh user** *user-name* **service-type { stelnet | all }** 例如：[HUAWEI] **ssh user** winda **service-type stelnet**	配置SSH用户的服务方式。命令中的参数和选项说明如下。 （1）*user-name*：指定前面创建的SSH用户账户名 （2）**stelnet**：二选一选项，指定*user-name*参数指定的SSH用户账户仅支持Stelnet服务 （3）**all**：二选一选项，指定*user-name*参数指定的SSH用户账户支持SFTP服务方式和STelnet服务方式 缺省情况下，SSH用户的服务方式是空，即不支持任何服务方式，可用**undo ssh user** *username* **service-type**命令取消指定SSH用户所有支持的服务方式
5	**ssh user** *user-name* **authorization-cmd aaa** 例如：[HUAWEI]**ssh user** winda **authorization-cmd aaa**	（可选）为指定的SSH用户配置按命令行授权。参数*user-name*就是前面创建的SSH用户。该命令只对使用RSA或DSA验证方式的SSH用户有效，授权后再进行相关AAA授权配置，否则该SSH用户的命令行授权配置不能生效。对于采用密码方式登录的用户，由AAA的配置决定是否需要授权。通常是不需要配置的 通常情况下，某级别用户经过授权后可以执行该级别及该级别以下的命令集。为了加强对用户权限的限制，实现权限的最小化控制，可以配置按命令行授权。配置按命令行授权后，用户输入的每一条命令都需要进行授权，授权通过后才可以执行，否则不能执行该命令。可通过**authorization-cmd** *privilege-level* **hwtacacs [local]**命令配置按命令行授权，生效后由参数*privilege-level*指定的级别用户执行命令时，每条命令都需要经过HWTACACS服务器授权

说明

这里的password验证要依靠AAA实现，所以当用户使用password、password-rsa 或password-dsa验证方式登录交换机时，需要在AAA视图下创建同名的本地用户。

如果SSH用户使用password验证，则只需要在SSH服务器端生成本地RSA或DSA密钥；如果SSH用户使用RSA或DSA验证，则在服务器端和客户端都需要生成本地RSA或DSA密钥对，并且服务器端和客户端都需要将对方的公钥配置到本地。

如果对SSH用户进行password验证（password、password-dsa或password-rsa验证）还需进行表 3-18 所示配置；如果对 SSH 用户进行 rsa 或 dsa 验证（包括 dsa、rsa、password-dsa或password-rsa验证）还需要进行表3-19所示配置；如果对SSH用户进行password-rsa或password-dsa验证，则AAA用户和RSA或DSA公钥都需要进行配置，即要同时进行表3-18和表3-19中的配置。

表3-18 配置对SSH用户进行password、password-dsa或password-rsa验证

步骤	命令	说明
1	**system-view** 例如：<HUAWEI> **system-view**	进入系统视图
2	**aaa** 例如：[HUAWEI] **aaa**	进入 AAA 视图
3	**local-user** *user-name* **password cipher** *password* 例如：[HUAWEI-aaa] **local-user** winda **password cipher**	配置本地用户名和密码。参数 *user-name* 的长度范围为 1～64 个字符，不支持空格，不区分大小写；参数 *password* 为长度范围为 1～16 个明文字符，或长度为 32 个密文字符，支持空格、单引号和问号，**区分大小写**
4	**local-user** *user-name* **service-type ssh** 例如：[HUAWEI-aaa] **local-user** winda **service-type ssh**	配置指定本地用户的服务方式为 SSH 服务
5	**local-user** *user-name* **privilege level** *level* 例如：[HUAWEI-aaa] **local-user** winda **privilege level** 10	配置指定本地用户为指定的级别，取值范围为 0～15 的整数

表3-19 配置对SSH用户进行dsa、rsa、password-dsa或password-rsa验证

步骤	命令	说明
1	**system-view** 例如：<HUAWEI> **system-view**	进入系统视图
2	**rsa peer-public-key** *key-name* [**encoding-type** { **der** | **openssh** | **pem** }] 例如：[HUAWEI] **rsa peer-public- key** 002 **encoding- type der** 或 **dsa peer-public-key** *key-name* **encoding-type** { **der** | **openssh** | **pem** } 例如：[HUAWEI] **dsa peer-public-key** 002 **encoding-type der**	配置 RSA 或 DSA 公共密钥编码格式，进入 RSA 或 DSA 公钥视图。命令中的参数和选项说明如下。 （1）*key-name*：指定 RSA 或 DSA 公钥名称，长度范围是 1～30 个字符，不支持大小写，不支持空格 （2）**encoding-type**：可选项，指定 RSA 或 DSA 公钥编码格式类型 （3）**der**：多选一选项，指定 RSA 或 DSA 公钥编码格式是 DER。DER 格式是将数据进行十六进制编码。此为缺省编码格式 （4）**openssh**：多选一选项，指定 RSA 或 DSA 公钥编码格式是 OpenSSH。OpenSSH 格式是将数据进行六十四进制编码

（续表）

步骤	命令	说明
2	**rsa peer-public-key** *key-name* [**encoding-type** { **der** | **openssh** | **pem** }] 例如：[HUAWEI] **rsa peer- public-key** 002 **encoding- type der** 或 **dsa peer-public-key** *key-name* **encoding-type** { **der** | **openssh** | **pem** } 例如：[HUAWEI] **dsa peer-public-key** 002 **encoding-type der**	（5）**pem**：多选一选项，指定 RSA 或 DSA 公钥编码格式是 PEM。PEM 格式是将数据进行六十四进制编码 【注意】命令中的 [**encoding-type** { **der** | **openssh** | **pem** }] 部分选项仅 S5700、S6700、S7700 和 S9700 系列交换机支持 可用 **undo rsa peer-public-key** *key-name* 命令删除指定的公钥 通过本命令指定 RSA 或 DSA 公钥编码格式后，交换机会自动生成相应编码格式的密钥，同时进入 RSA 或 DSA 公钥视图，再执行 **public-key-code begin** 命令后，用户即可通过手动复制的方式将客户端产生的公钥复制到服务器端。客户端的公钥是由客户端软件随机生成的
3	**public-key-code begin** 例如：[HUAWEI-rsa-public-key]**public-key-code begin** 或 [HUAWEI-dsa-public-key] **public-key-code begin**	进入公钥编辑视图。输入本命令后，进入公钥编辑视图，在该视图下可以开始输入密钥数据。**这里的公钥就是在表 3-16 中第 2 步所创建的本地密钥对中的公钥**。在输入密钥数据时，字符之间可以有空格，也可以按回车键继续输入数据。所配置的公钥必须是按公钥格式编码的十六进制字符串，是由支持 SSH 的客户端软件随机生成的，具体操作参见相应的 SSH 客户端软件的帮助文档
4	**public-key-code end** 例如：[HUAWEI-rsa-key-code]**public-key-code end** 或 [HUAWEI-dsa-key-code]**public-key-code end**	从公钥编辑视图退回到公钥视图，并且保存用户配置的公钥。如果未输入合法的密钥编码数据，执行本步骤后也将无法生成密钥 如果指定的密钥 *key-name* 已经在别的窗口下被删除，再执行本步骤时，系统会提示：密钥已经不存在，此时直接退到系统视图
5	**peer-public-key end** 例如：[HUAWEI-rsa-public-key]**peer-public-key end** 或 [HUAWEI-dsa-public-key] **peer-public-key end**	退出公钥视图，回到系统视图
6	**ssh user** *user-name* **assign** { **rsa-key** | **dsa-key** } *key-name* 例如：[HUAWEI] **ssh user** winda **assign rsa-key** 002	为用户分配一个已经存在的公钥。命令中的参数和选项说明如下： （1）*user-name*：指定 AAA 定义的有效 SSH 用户名 （2）**rsa-key**：二选一选项，指定使用 RSA 公钥 （3）**dsa-key**：二选一选项，指定使用 DA 公钥（**S2700 和 S3700 系列不支持**） （4）*key-name*：指定配置的客户端 RSA 或者 DSA 公钥名，要与第 2 步指定的公钥名一致 缺省情况下，没有为 SSH 用户分配公钥，可用 **undo ssh user** *user-name* **assign** { **rsa-key** | **dsa-key** }命令删除指定用户和对应公钥之间的对应关系

【示例 1】进入RSA公钥视图。

```
<HUAWEI>system-view
[HUAWEI] rsa peer-public-key 002
Enter "RSA public key" view, return system view with "peer-public-key end".
[HUAWEI-rsa-public-key]
```

【示例 2】进入公钥编辑视图，输入RSA密钥。

```
[HUAWEI] rsa peer-public-key 003
Enter "RSA public key" view, return system view with "peer-public-key end".
[HUAWEI-rsa-public-key] public-key-code begin
Enter "RSA key code" view, return last view with "public-key-code end".
[HUAWEI-rsa-key-code] 308186028180739A291ABDA704F5D93DC8FDF84C427463
[HUAWEI-rsa-key-code] 1991C164B0DF178C55FA833591C7D47D5381D09CE82913
[HUAWEI-rsa-key-code] D7EDF9C08511D83CA4ED2B30B809808EB0D1F52D045DE4
[HUAWEI-rsa-key-code] 0861B74A0E135523CCD74CAC61F8E58C452B2F3F2DA0DC
[HUAWEI-rsa-key-code] C48E3306367FE187BDD944018B3B69F3CBB0A573202C16
[HUAWEI-rsa-key-code] BB2FC1ACF3EC8F828D55A36F1CDDC4BB45504F020125
```

[HUAWEI-rsa-key-code] public-key-code end

[HUAWEI-rsa-public-key]

【示例 3】为用户zhangsan分配RSA公钥key1。

<HUAWEI>system-view

[HUAWEI] ssh user zhangsan assign rsa-key key1

【示例 4】为130.100.0.114用户分配DSA公钥pemkey。

<HUAWEI>system-view

[HUAWEI] ssh user 130.100.0.114 assign dsa-key pemkey

Info: Succeeded in adding a new SSH user.

【示例 5】配置对用户john按命令行授权。执行命令后会有提示信息，要求已为该用户配置了命令行授权方法。可通过authorization-cmdprivilege-levelhwtacacs [local] 命令配置。

<HUAWEI>system-view

[HUAWEI] ssh user john authorization-cmd aaa

Info: Please make sure that the command line authorization method has been set for the user.

4．用户通过STelnet登录交换机

完成以上各部分的配置后，用户就可以使用安装了SSH客户端软件的终端通过Stelnet方式登录到交换机。此处以使用第三方 SSH 客户端软件putty，采用比较简单的密码验证方式为例介绍通过Stelnet登录到交换机的操作方法。

（1）打开 putty 软件配置对话框，在“Host Name(or IP adress)”文本框中输入交换机的主机名或 IP 地址；在“port”下拉列表中选择所配置的SSH服务端口，缺省为22；在“Protocol”栏中选择“SSH”单选项，如图3-10所示。

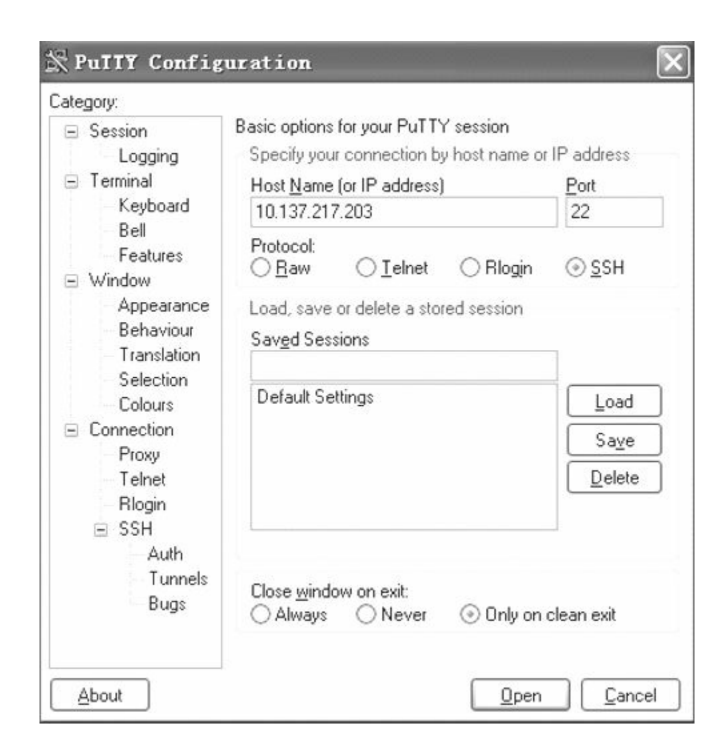


图3-10 putty配置对话框

（2）其他选择采用缺省配置即可，然后单击“Open”按钮即可登录到交换机。首先在 putty 终 端软件控制台中出现如下提示信息，请你输入用户名和密码（输入的密码不会在控制台显示的）。完成后，如果配置正确，就可以成功登录到交换机了。

login as: client001　#---输入登录用户名

Sent username "client001"

client001@10.137.217.203's password:　#---输入用户密码（不会在控制台中显示的）

Info: The max number of VTY users is 8, and the number

of current VTY users on line is 5.

The current login time is 2012-08-06 09:35:28.

<SSH Server>

5．STelnet登录管理

Stelnet成功登录后，可以执行以下任意视图display命令查看相关信息。

（1）使用display ssh user-information [username] 命令在SSH服务器端（也就是交换机）查看SSH用户配置信息。如果不指定SSH用户，则可以查看SSH服务器端所有的SSH用户配置信息。

（2）使用display ssh server status命令查看SSH服务器的全局配置信息。

（3）使用display ssh server session命令在SSH服务器端查看与SSH客户端连接的会话信息。

3.6.5 通过STelnet登录交换机的配置示例

本配置示例拓扑结构如图3-11所示。终端PC1、PC2和担当SSH服务器的S系列交换机之间路由可达，10.137.217.203是SSH服务器的管理口IP地址。在SSH服务器端配置两个登录用户为client001和client002，PC1使用cliet001用户通过password验证方式登录SSH服务器，PC2使用cliet002用户通过RSA验证方式登录SSH服务器。

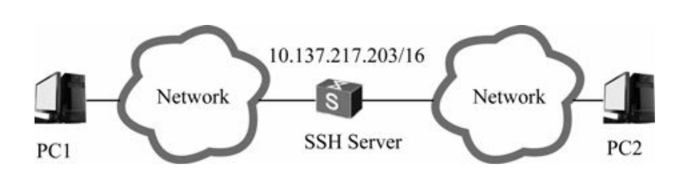


图3-11 通过STelnet登录交换机的配置示例拓扑结构

1．基本配置思路

本示例要求采用STelent方式（采用的是SSH服务）登录到交换机VRP系统，所使用的也是VRP系统的VTY用户界面，但要求在登录终端安装SSH服务客户端软件。本示例中最关键是的要求其中一个用户采用password验证方式登录，另一个用户采用RSA验证方式登录。根据3.6.3节介绍的配置任务，可以得出本示例的基本配置思路如下。

（1）因为PC1终端用户采用的是password验证方式登录到交换机，所以需要事先安装好SSH 服务客户端软件；PC2 终端采用的是RSA 验证方式登录到交换机，除了需要安装SSH服务客户端软件外，还需要生成用于RSA验证所需的本地RSA公钥对和服务器密钥对。

说明

SSH用户有password、RSA、password-rsa、DSA、password-dsa和all这6种认 证方式。如果SSH用户的认证方式为password、password-rsa、password-dsa时，必须在服务器端配置同名的本地用户；如果SSH用户的认证方式为RSA、password-rsa、DSA、password-dsa和all，则在服务器端应保存SSH客户端的RSA或DSA

公钥。

（2）配置STelnet登录所用的VTY用户界面，并设置它们支持SSH服务，采用AAA验证方式和用户级别。

（3）在配置为SSH服务器的交换机端生成本地密钥对和服务器密钥对，实现在SSH服务器端和SSH客户端进行安全的数据交互。

（4）在SSH服务器端开启STelnet服务功能，并创建SSH用户client001和client002，分别指定password验证方式和RSA验证方式，配置client001用户密码、用户级别和支持SSH服务。

（5）用户client001和client002分别以STelnet方式实现登录SSH服务器。

2．具体配置步骤

根据以上配置思路，再结合3.6.4节介绍的配置任务，可以得出如下具体配置步骤。

（1）配置STelnet登录所用的VTY用户界面属性，包括指定AAA验证方式，支持SSH服务和用户级别。

```
<HUAWEI> system-view
[HUAWEI] sysnameSSH Server
[SSH Server] user-interface vty 0 4
[SSH Server-ui-vty0-4] authentication-mode aaa
[SSH Server-ui-vty0-4] protocol inbound ssh
[SSH Server-ui-vty0-4] user privilege level 5
[SSH Server-ui-vty0-4] quit
```

（2）新建用户名为 client001 的 SSH 用户，且验证方式为 password，并为其配置AAA验证所需的密码（本示例为huawei@123），用户级别（本示例中为3级）和SSH服务支持。

```
[SSH Server] ssh user client001 authentication-type password
[SSH Server] aaa
[SSH Server-aaa] local-user client001 password cipherhuawei@123
[SSH Server-aaa] local-user client001 privilege level 3
[SSH Server-aaa] local-user client001 service-type ssh
[SSH Server-aaa] quit
```

（3）新建用户名为client002的SSH用户，且验证方式为RSA。

```
[SSH Server] ssh user client002 authentication-type rsa
```

（4）在PC1和PC2上分别安装支持SSH服务的Putty终端软件。然后在PC2上运行puttygen.exe程序，打开如图3-12所示对话框，生成公钥和私钥两个文件，供后续使用。

（5）选择“SSH2 RSA”单选项（采用缺省的 1 024位密钥），然后单击“Generate”按钮，进入密钥生成状态，此时只要将鼠标在空白处进行移动，就可以看到生成的RSA密钥，如图 3-13所示。单击“Save public key”按钮保存公钥文件名为 key.pub；单击“Save private key”按钮保存私钥文件名为 private.ppk，在弹出的提示对话框中单击“Yes”按钮即可。这样就生成了公钥/私钥对。

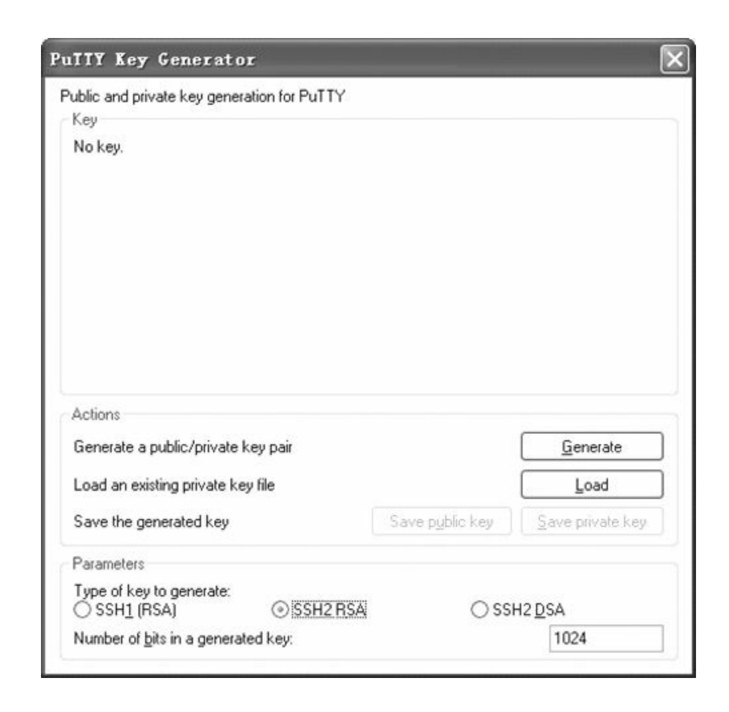


图3-12 密钥生成对话框

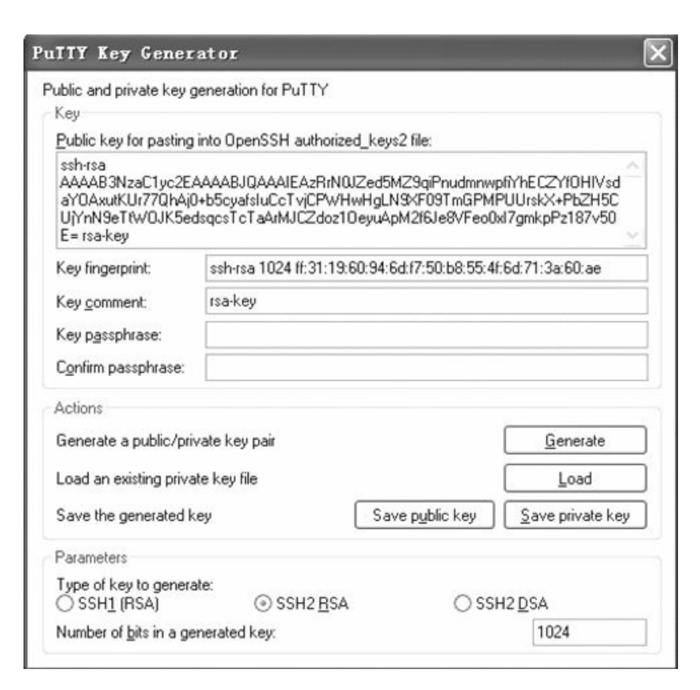


图3-13 生成密钥后的密钥生成对话框

（6）再在PC2上运行sshkey.exe程序，打开如图3-14所示对话框，单击“Browser”按钮，在找到的对话框中打开上一步保存的公有密钥文件key.pub。然后单击“Convert(C)”按钮，打开如图 3-15所示对话框，即可在“RSA public-key after convert”栏中见到转换后的RSA公钥。复制其中的内容，以文件形式保存，以备后用。

ssh key convert

SSH Public-key file name:

Browse(B)

Exit(X)

C:\SSH\key.pub

RSA public-key after convert

Convert(C)

Save(S)

RSA public-key before convert

图3-14 SSH密钥转换对话框



图3-15 生成公钥后的SSH密钥转换对话框

（7）在 SSH服务器端使用 rsa local-key-pair create命令生成本地RSA密钥对（密钥位也为 1 024位），用于服务器端向 SSH用户 client002传输数据时的数据加密保护。

[SSH Server] rsa local-key-pair create

The key name will be: SSH Server_Host

The range of public key size is (512 ～ 2048).

NOTES: If the key modulus is greater than 512,

it will take a few minutes.

Input the bits in the modulus[default = 2048] :1024

Generating keys. .

...........++++++++

...........................++++++++

....+++++++++

...+++++++++

（8）在SSH服务器端配置PC2端上产生的RSA公钥，并为SSH用户创建client002绑定在PC2上创建的RSA公钥，以实现SSH服务器对SSH用户client002的验证。在执行public-key-code begin命令后的提示符下输入前面在第2步中得到的 client002客户端RSA公钥。

[SSH Server] rsa peer-public-key rsakey001

Enter "RSA public key" view, return system view with "peer-public-key end".

[SSH Server-rsa-public-key] public-key-code begin

Enter "RSA key code" view, return last view with "public-key-code end".

[SSH Server-rsa-key-code] 30818702 818100CD 1ACDD096 5E779319 F6A88F9E E7669F0A

[SSH Server-rsa-key-code] 5F898844 09961F38 7215B1D6 98380C6E B4A52BEF B421023D

[SSH Server-rsa-key-code] 3E6F9732 69FB08B8 2713BE30 8F587C07 80B37D5C 5D3D4E61

[SSH Server-rsa-key-code] 8F30F514 AEC917F8 F6D91F90 948D89CD F5E4ED58 E24AE5E7

[SSH Server-rsa-key-code] 6CA9CB13 713680AC C24265DA 33D4E7B2 B80A4CD9 FE897BC5

[SSH Server-rsa-key-code] 457A8D31 23B82692 93F3D7CE EFE74102 0125

[SSH Server-rsa-key-code] public-key-code end

[SSH Server-rsa-public-key] peer-public-key end

[SSH Server] ssh user client002 assign rsa-key rsakey001 #---把以上创建的RSA密钥与 client002客户进行绑定

（9）在SSH服务器上使能STelnet服务功能，并配置SSH用户client001、client002的服务方式为Stelnet。

[SSH Server] stelnet server enable

[SSH Server] ssh user client001 service-type stelnet

[SSH Server] ssh user client002 service-type stelnet

（10）通过STelnet登录交换机。

在PC1端的client001用户打开putty软件，输入交换机的IP地址，选择协议类型为SSH（如图3-16所示），用password验证方式连接SSH服务器。单击“Open”按钮即在putty 终端界面出现如下提示，然后正确输入前面所配置的 client001 用户名和密码，按回车键即可成功登录到交换机。

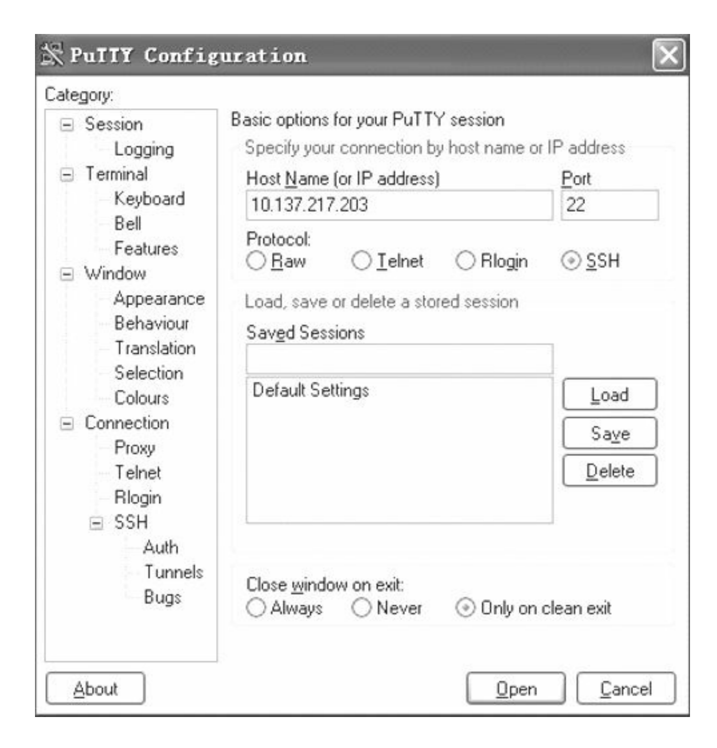


图3-16 PC1上的putty Stelnet登录连接配置对话框

login as: client001

Sent username "client001"

client001@10.137.217.203's password:

Info: The max number of VTY users is 8, and the number

of current VTY users on line is 5.

The current login time is 2012-08-06 09:35:28.

<SSH Server>

在PC2端的client002用户采用RSA验证方式连接SSH服务器。首先也要打开putty软件配置对话框，输入交换机的IP地址，选择协议类型为SSH，参见图3-16。然后单击左侧导航栏“Connection→SSH”，出现如图3-17所示对话框，在“Preferred SSH protocol version”栏中选择“2”（即SSH协议版本2）单选项。再单击左侧导航栏“Connection→SSH”下面的“Auth”（验证），打开如图3-18所示对话框。

private.ppk。最后在图3-18所示对话框中单击“Open”按钮，在控制台中出现的提示信息下输入client002用户名，按回车键后即可成功登录到交换机。

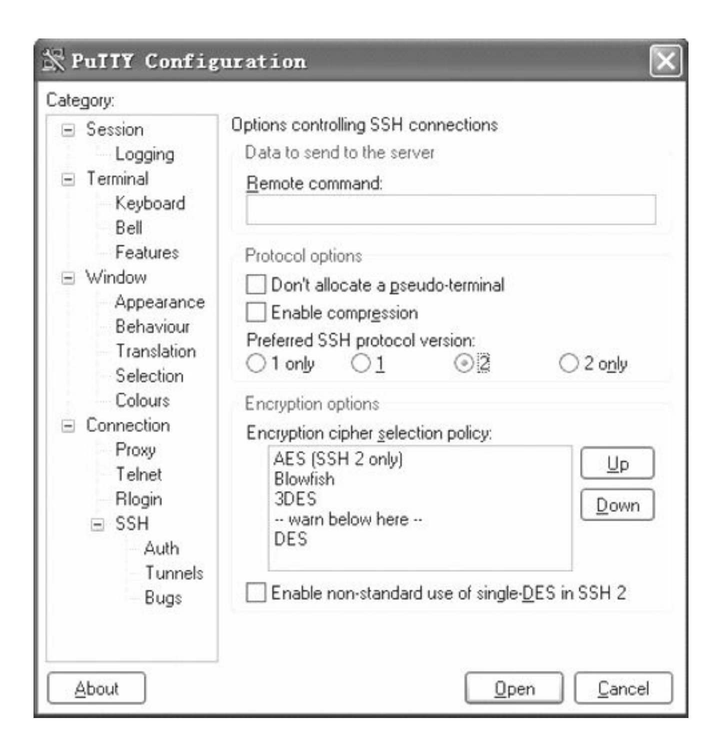


图3-17 SSH服务器连接配置对话框

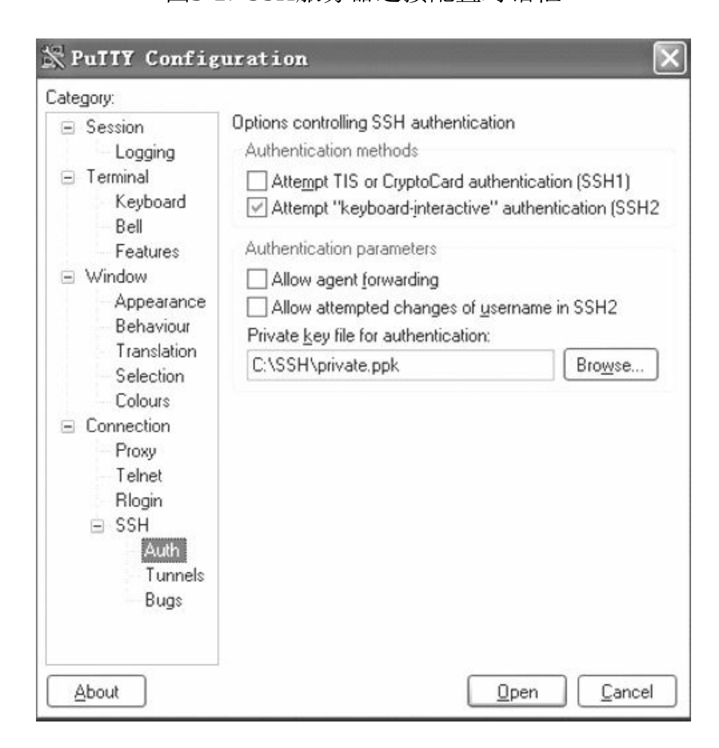


图3-18 SSH验证配置对话框

在“Private Key file for authentication”文本框中选择与配置到服务器端的RSA公钥对应的私钥文件private.ppk。最后在图3-18所示对话框中单击“Open”按钮，在控制台中出现的提示信息下输入client002用户名，按回车键后即可成功登录到交换机。

login as: client002

Authenticating with public key "rsa-key"

Info: The max number of VTY users is 8, and the number

of current VTY users on line is 5.

The current login time is 2012-08-06 09:35:28.

<SSH Server>

3.6.6 配置用户通过HTTP Web网管登录交换机

在配置用户通过HTTP登录交换机之前，也需要确保终端与交换机之间路由可达。整个HTTP登录方式的配置任务如下。

（1）上传和加载Web网页文件：在使能HTTP服务功能前，需要确保交换机上已经加载了Web网页文件。

（2）配置SSL策略并加载数字证书：但仅在需重新加载**SSL**证书时才执行此项配置任务 。

（3）配置HTTP服务功能：包括HTTPS及HTTP服务的使能、端口号、会话超时时间等。

（4）配置HTTP用户：包括HTTP用户名及密码、用户级别、接入类型等。这部分与前面的Console用户界面中的AAA验证用户配置方法差不多，具体参见3.4.4节表3-8中的第 6～8步，不同的只是在第 8步中要使用 local-user user-name service-type http命令配置对应用户对HTTP服务的支持。

（5）配置HTTP访问控制：包括配置ACL规则及HTTP基本访问控制列表，提高HTTP访问的安全性。仅在需要**ACL**控制用户通过**HTTP**方式登录交换机时才执行此项配置任务 。

（6）用户通过HTTP登录交换机。

与HTTP登录交换机相关的参数缺省配置如表3-20所示。

表3-20 HTTP登录交换机的相关参数缺省值

参数	缺省值
SSL 策略	有缺省 SSL 策略
HTTP 服务功能	未使能
HTTPS 服务功能	未使能
HTTP 服务器监听端口号	80
HTTP 会话超时时间	20min
HTTP 用户	用户名：admin，密码：admin

下面介绍以上配置中除第4项外的其他五项配置任务的配置方法。

1．上传和加载Web网页文件

华为S系列交换机出厂时在存储交换机中就已经保存了Web网页文件，首次使用时不需要上传Web网页文件（但仍需要进行加载操作 ）。当需要将交换机从当前版本升级至更高版本时，必须重新上传Web网页文件。Web网页文件获取路径：请先登录华为公司企业业务支持网站（http://support.huawei.com/enterprise），登录后，在“软件下载 >产品软件 >企业网络 >交换机 >园区交换机”路径下根据产品型号和版本名称，下载对应的版本软件。版本软件中包含Web网页文件，名称为“产品-软件版本号.WEB网管文件版本号.web.7z”。

可通过FTP、SFTP等方式上传Web网页文件，具体请参见本章后面3.7节介绍的远程文件管理方法。加载Web网页文件的方法很简单，只需在系统视图下通过http server load file-name命令加载指定的Web网页文件即可。

Web网页文件中包含SSL证书，用HTTP方式登录时进行SSL验证，确保用户信息的安全（当前HTTP登

录会跳转至HTTPS登录，登录成功后跳转回HTTP）。此证书也可以用于HTTPS方式的登录，同时确保用户信息及交互数据的安全。另外，用户可以重新加载新的数字证书。交换机重启时，如果重启前加载的Web网页文件不存在，HTTP 服务功能将不能使能。如果要取消当前加载的文件，必须先加载另外一个文件，否则无法被取消 。

Web网页文件必须保存在存储器根目录下 ，且必须是“*.web.zip”或“*.web.7z”格式，为 4～64 个字符，不支持空格。缺省情况下，使能 HTTP 服务功能系统缺省加载名称为*.web.7z的网页文件，可用undo http server load命令加载系统缺省的Web网页文件。如果要使用的Web网页文件名不是*web.7z，则需要执行http server load命令重新加载。

【示例 1】加载文件名为web_1.web.7z的Web网页文件。

<HUAWEI>system-view

[HUAWEI] http server loadweb_1.web.7z

2．（可选）配置SSL策略并加载数字证书

通过HTTP登录时也会跳转至通过HTTPS登录（以保证登录时用户信息的安全），所以需要在交换机上为HTTP服务配置SSL策略。交换机上提供缺省的SSL策略，同时Web网页文件中也包含SSL证书，所以用户可以不必再上传证书以及手动配置SSL策略。当然，为了保证安全性也可以从CA（Certificate Authority）处重新获取有效证书，然后进行手动配置SSL策略。

这个步骤的配置方法请参见3.6.8节介绍的通过HTTPS Web网管方式登录交换机中的“上传服务器数字证书文件及私钥文件”（必须保存在**flash**：存储器根目录的**security**目录中 ）和“配置SSL策略并加载数字证书文件”这两部分。

当用户手动为HTTP服务器配置了SSL策略后，则以用户配置的SSL策略生效。

3．配置HTTP服务功能

此处配置的是HTTP服务属性和基本的管理操作，包括启用HTTP、HTTPS服务功能，配置SSL策略、HTTP服务端口、HTTP会话超时和释放HTTP连接。具体配置步骤如表3-21所示。

表3-21 HTTP服务功能的配置步骤

步骤	命令	说明
1	**system-view** 例如：<HUAWEI> **system-view**	进入系统视图
2	**http secure-server ssl-policy** *policy-name* 例如：[HUAWEI]**http secure-server ssl-policy** http_server	（可选）为服务器创建SSL策略。策略名的长度范围为1～23个字符，不支持空格，支持“_”、字母和数字，不区分大小写。缺省情况下，交换机提供缺省的SSL策略（Default），当加载Web网页文件后，会自动加载缺省的SSL策略，而无需手动配置，**此步仅需在重新配置了SSL策略的情况下执行**，可用**undo http secure-server ssl-policy**命令恢复HTTP服务器的SSL策略为缺省策略
3	**http secure-server enable** 例如：[HUAWEI]**http secure- server enable**	使能HTTPS服务功能。在**使能HTTP服务功能前，必须先使能HTTPS服务**。执行本命令使能HTTP安全服务功能后，用户必须经过认证后才能通过浏览器输入网址访问Web网管系统进行交换机管理 如果Web网页文件未被加载或加载不成功，使能HTTPS服务会提示失败；如果没有执行上个步骤（配置SSL策略），系统会自动加载缺省的SSL策略。但此时必须保证已加载了正确的包含证书的Web网页文件 缺省情况下，HTTPS服务功能未使能，**undo http secure-server enable**命令去使能HTTPS服务功能
4	**http server enable** 例如：[HUAWEI] **http server enable**	使能HTTP服务功能。执行本命令使能HTTP服务功能后，可以通过浏览器输入交换机的IP地址访问Web网管系统管理交换机 缺省情况下，HTTP服务功能处于去使能状态，可用**undo http server enable**命令用来去使能HTTP服务功能
5	**http server port** *port-number* 例如：[HUAWEI] **http server port** 8080	（可选）配置HTTP服务器监听端口号，取值范围为80或1 025～55 535。配置监听端口号，可以有效防止攻击者对HTTP服务标准端口的访问，增加交换机的安全性 缺省情况下，HTTP服务器监听端口号是80，可用**undo http server port**命令恢复HTTP服务器监听的端口号到缺省值
6	**http timeout** *timeout* 例如：[HUAWEI] **http timeout** 15	（可选）配置HTTP会话的超时时间（也即闲置时间），取值范围为1～60的整数分钟。执行本命令后，所有登录系统的Web用户的超时时间都相同。如果用户超时，用户将自动下线，HTTP服务器不会主动通知用户，而是等待用户发送下一次请求时再通知用户 本命令是覆盖式命令，以最后一次配置为准。缺省情况下，会话超时时间为20min，可用**undo http timeout**命令恢复HTTP服务器的超时时间为缺省值
7	**free http user-id** *user-id* 例如：[HUAWEI] **free http user-id** 90	（可选）释放指定Web界面编号的HTTP用户，取值范围为1～256的整数。目前，交换机只支持登录5个HTTP用户，可通过此命令可以手动释放Web用户界面 **S2700和S3700系列交换机不支持本命令**

4．（可选）配置HTTP访问控制

与通过Console用户界面和VTY用户界面登录交换机一样，Web网管登录方式也可以通过ACL来控制。用户可以通过基本ACL允许指定的客户端通过HTTP方式登录到交换机，以提高安全性。一般不需要配置。

当ACL中的规则选择permit选项时，则允许指定源IP地址的其他交换机与本交换机建立HTTP连接；当ACL的规则选择deny选项时，则拒绝指定源IP地址的其他交换机与本交换机建立HTTP连接；当ACL未配置规则时，则允许任何其他交换机与本交换机建立HTTP连接。具体配置步骤如表3-22所示。

表3-22 HTTP访问控制的配置步骤

步骤	命令	说明
1	**system-view** 例如：<HUAWEI> **system-view**	进入系统视图
2	**acl** [**number**] *acl-number* 例如：[HUAWEI] **acl** 2000	进入 ACL 视图。HTTP 只支持基本访问控制列表，列表号为 2000～2999
3	**rule** [*rule-id*] { **deny** | **permit** } [**source** {*source-address source-wildcard* | **any** } | **fragment**| **logging** | **time-range** *time-name*] * 例如：[HUAWEI-acl-basic-2000]**rule** 5 **permit source** 192.168.32.1 0	配置 ACL 规则。具体参数说明参见本书第 9 章
4	**quit** 例如：[HUAWEI-acl-basic-2000]**quit**	退回到系统视图
5	**http acl** *acl-number* 例如：[HUAWEI] **http acl** 2000	配置通过基本 ACL 控制用户通过 HTTP 方式访问交换机。如果 ACL 中没有配置规则，则 HTTP 服务器将不会拒绝用户登录，如果存在已登录并符合 ACL 规则中过滤条件的用户，HTTP 服务器不会主动将客户端踢下线，而是等待客户端发送下一次请求时，再按照配置的 ACL 规则过滤用户 重复执行本命令，新配置覆盖旧配置。缺省情况下，没有为 HTTP 服务器配置 ACL，可用 **undo http acl** 命令删除 HTTP 服务器的 ACL

5．用户通过HTTP登录交换机

完成以上配置后，就可以在PC上打开Web浏览器，在地址栏中直接输入http://IP （要确保 PC 和交换机之间有可达的路由），按回车键后将显示如图 3-19 所示的 Web登录对话框。分别输入之前设置的 Web 网管账号和密码，输入验证码，并选择 Web网管系统的语言后从地址栏中就可看到当前的登录页面已跳转到 HTTPS 的登录页面。单击“登录”按钮或直接按回车键即可进入 Web 网管系统主页面。此时从地址栏中可以看到，页面又跳转回到了 HTTP 页面。登录到 Web 网管后，可以对交换机进行管理和维护。

6．HTTP Web网管登录管理

登录成功后，可在交换机 VRP 命令行界面执行 display http user [username username] 命令查看当前在线用户的摘要信息或指定用户的详细信息；执行display http server命令查看当前HTTP服务器信息。



图3-19 HTTP Web网管登录界面

3.6.7 通过HTTP Web网管登录交换机的配置示例

本示例拓扑结构如图3-20所示，现要从PC上通过HTTP方式登录到交换机上，将交换机作为Web网管服务器，实现图形化界面管理和维护交换机。

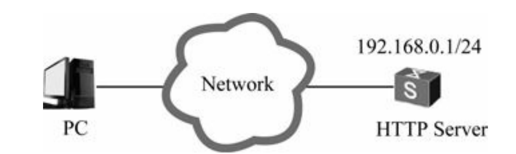


图3-20 通过HTTP登录交换机的配置示例拓扑结构

1．配置思路

HTTP Web网管登录方式的配置比较简单，根据 3.6.6节介绍的配置任务，以及本示例的要求可以得出如下基本配置思路。

（1）向交换机上传并加载Web网页文件。

（2）同时使能HTTPS和HTTP服务功能，配置HTTP服务属性参数，如HTTP服务监听端口和超时时间。当然也可以不配置这些属性参数，因为它们都有缺省值。

（3）在AAA视图下创建用于HTTP Web网管登录的本地用户账户，配置用户级别和对HTTP服务的支持。

（4）在浏览器地址栏中输入交换机的管理IP地址即可实现成功登录。

2．具体配置步骤

（1）上传并加载Web 网页文件。交换机存储器中已配置Web 文件，仅当需要更新Web网页才需要重新上传和加载。文件上传方法可以通过FTP、SFTP、SCP等传输协议进行，具体将在本章后面3.7节介绍。

Web网页文件上传后可以通过dir用户视图命令查看，如下所示可以得知交换机中的Web文件名为webtest.7z（粗体字部分显示）。

```
<HTTP-Server>dir
Directory of flash:/
  Idx   Attr   Size(Byte)   Date   Time   FileName
  0   -rw-   524,558   Apr 14 2011 16:24:39   private-data.txt
  1   -rw-   1,302   Apr 14 2011 19:22:30   back_time_a
  2   -rw-   951   Apr 14 2011 19:22:35   back_time_b
  3   drw-   -   Apr 09 2011 19:46:14   src
  4   -rw-   421   Apr 09 2011 19:46:14   vrpcfg.zip
  5   -rw-   1,308,478   Apr 14 2011 19:22:45   webtest.7z
  6   drw-   -   Apr 10 2011 01:35:54   logfile
  7   -rw-   4   Apr 14 2011 04:56:35   snmpnotilog.txt
  8   drw-   -   Apr 11 2011 16:18:53   security
  9   drw-   -   Apr 13 2011 11:37:40   lam
..
65,233 KB total (7,289 KB free)
```

下面通过http server load命令在交换机上加载这个已上传的Web网页文件。

```
<HTTP-Server>system-view
[HTTP-Server] http server loadwebtest.7z
```

（2）使能HTTPS和HTTP服务功能。

```
[HTTP-Server] http secure-server enable
[HTTP-Server] http server enable
```

（3）创建HTTP用户（此处的用户名为admin，密码为huawei），并配置用户级别（此处为最高的15级）和对HTTP服务的支持。

[HTTP-Server] aaa

[HTTP-Server-aaa] local-user admin password cipher huawei

[HTTP-Server-aaa] local-user admin privilege level 15

[HTTP-Server-aaa] local-user admin service-type http

[HTTP-Server-aaa] quit

（4）通过HTTP协议登录交换机。

在用户PC的Web浏览器地址栏中直接输入http://192.168.0.1（此处假设交换机的管理IP地址为192.168.0.1），按回车键后，将显示如图3-19所示的登录对话框。正确输入HTTP用户名、密码和验证码，单击“登录”按钮或直接按回车键即可进入交换机的Web网管系统主页面。

在交换机的命令行界面下执行display http server任意视图命令可以看到交换机配置的HTTP服务器的当前状态，可以验证配置是否正确。如下所示：

[HTTP-Server] display http server

HTTP Server Status :enabled

HTTP Server Port : 80(80)

HTTP Timeout Interval : 20

Current Online Users : 1

Maximum Users Allowed : 5

HTTP Secure-server Status : enabled

HTTP Secure-server Port : 443(443)

HTTP SSL Policy :Default

3.6.8 配置用户通过HTTPS Web网管方式登录交换机

HTTPS Web网管方式比普通的HTTP Web网管方式更加安全，因为HTTPS方式将HTTP和SSL结合，通过SSL对服务器身份进行验证，对传输的数据进行加密，从而实现了对交换机的安全管理。在配置用户通过HTTPS 登录交换机之前，也需要确保终端 与交换机之间路由可达。整个HTTPS登录方式的配置任务如下。

（1）上传及加载Web网页文件：参见3.6.6节第1点。

（2）上传服务器数字证书文件及私钥文件：通过文件上传方式将数字证书文件和私钥文件上传至交换机。仅在需重新加载**SSL**证书时才执行此项配置任务 。

（3）配置 SSL策略并加载数字证书：仅在需重新加载 **SSL** 证书时才执行此项配置任务 。

（4）配置HTTPS服务功能：包括HTTPS服务的使能、端口号、会话超时时间等。具体配置方法与3.6.6节第3点介绍的HTTP服务功能配置方法差不多。

（5）配置HTTP用户：包括HTTP用户名及密码、用户级别、接入类型等。这部分也与前面的Console用户界面中的AAA验证用户配置方法差不多，具体参见3.4.4节表3-8中的第6～8步，不同的只是在第 8步中要使用 local-useruser-name service-type http命令配置对应用户对HTTP服务的支持。

（6）配置HTTP访问控制：包括配置ACL规则及HTTP基本访问控制列表，提高访问的安全性。仅在需要**ACL**控制用户通过**HTTP**方式登录交换机时才执行此项配置任务 。参见3.6.6节第4点。

（7）用户通过HTTPS登录交换机：通过HTTPS方式登录交换机。

与HTTPS登录相关的参数缺省配置如表3-23所示。

表3-23 HTTPS登录交换机的相关参数缺省值

参数	缺省值
SSL 策略	有缺省的 SSL 策略
HTTPS 服务功能	未使能
HTTPS 服务器监听端口号	443
HTTP 会话超时时间	20min
HTTP 用户	用户名：admin，密码：admin

下面仅介绍以上第2、3、4、7项配置任务。

1．（可选）上传服务器数字证书文件及私钥文件

交换机上提供缺省的SSL策略，同时Web网页文件中也包含SSL证书，所以用户可以不必再上传证书以及手动配置SSL策略。但为了保证安全性，可以从CA处重新获取有效证书，然后进行手动配置SSL策略。

可使用FTP、SFTP或SCP方式将服务器数字证书和私钥文件上传至交换机（具体上传方法参数本章后面介绍的 3.7 节相关内容），且必须保存在 **flash**：存储器根目录的**security**目录中，如交换机无此目录，可执行mkdir security命令创建。

证书格式分为PEM格式、ASN1格式和PFX格式。虽然证书的格式不相同，但是证书的内容一样。PEM格式的证书是最常用的一种数字证书格式，文件的扩展名是.pem，适用于系统之间的文本模式传输。ASN1是通用的数字证书格式之一，文件的扩展名是.der，是大多数浏览器的缺省格式。PFX 是通用的数字证书格式之一，文件的扩展名是.pfx，是可移植的格式及二进制格式。

2．（可选）配置SSL策略并加载数字证书文件

加载数字证书文件的同时要指定私钥文件，具体的配置步骤如表3-24所示。

表3-24 SSL策略并加载数字证书的配置步骤

<table>
<tr><th>步骤</th><th>命令</th><th colspan="2">说明</th></tr>
<tr><td>1</td><td>**system-view**
例如：<HUAWEI> **system-view**</td><td colspan="2">进入系统视图</td></tr>
<tr><td>2</td><td>**ssl policy** *policy-name*
例如：[HUAWEI] **ssl policy** https_der</td><td colspan="2">创建 SSL 策略并进入 SSL 策略视图，策略名的长度范围是 1～23 个字符，不支持空格，支持“_”、字母和数字，不区分大小写</td></tr>
<tr><td rowspan="4">3</td><td>**certificate load pem-cert** *cert-filename* **key-pair** ｛**dsa** | **rsa**｝**key-file** *key-filename***auth-code cipher** *auth-code*
例如：[HUAWEI-ssl-policy-https_der] **certificate load pem-cert servercert. pem key-pair dsa key-file** serverkey. pem **auth-code cipher** 123456</td><td>加载PEM格式的证书</td><td rowspan="4">根据证书类型，选择其中一个命令。命令中的参数和选项说明如下。
（1）*cert-filename*：指定证书文件名称，为 1～64 个字符。该文件名由上传的文件决定，必须与上传的文件的文件名称一致，且必须保存在系统根目录下名为 **security** 的子目录下，如果没有 **security** 目录，则需要创建此目录
（2）**dsa**：二选一选项，指定密钥对类型是 DSA
（3）**rsa**：二选一选项，指定密钥对类型是 RSA
（4）*key-filename*：指定密钥文件名称，为 1～64 个字符。该文件名由上传的文件决定，必须与上传文件的文件名称一致，且必须保存在系统根目录下名为 **security** 的子目录下，如果没有 **security** 目录，则需要创建此目录
（5）*auth-code*：指定密钥文件验证码。密钥文件验证码用来进行身份验证，保证合法客户端安全登录服务器。输入明文密码时为 1～31 个字符；输入密文密码时为 32 或 56 个字符，区分大小写，不支持空格
（6）*mac-code*：指定消息验证码。消息验证码用来保证报文内容的完整性，即报文内容不被篡改，输入明文密码时为 1～31 个字符；输入密文密码时为 32 或 56 个字符，区分大小写，不支持空格
一个 SSL 策略只能加载一个证书或者证书链。如果已经加载了证书或者证书链，加载新证书或者证书链之前必须先卸载旧证书或者证书链</td></tr>
<tr><td>**certificate load asn1-cert** *cert-filename* **key-pair** ｛**dsa** | **rsa**｝**key-file** *key-filename*
例如：[HUAWEI-ssl-policy-https_der] **certificate load asn1-cert** servercert. der **key-pair rsa key-file** serverkey.der</td><td>加载ASN1格式的证书</td></tr>
<tr><td>**certificate load pfx-cert** *cert-filename* **key-pair** ｛**dsa** | **rsa**｝｛**mac cipher** *mac-code* |**key-file** *key-filename*｝**auth-code cipher** *auth- code*
例如：[HUAWEI-ssl-policy-https_der] **certificate load pfx-cert servercert. pfx key-pair rsa key-file** serverkey. pfx **auth-code** cipher %$%$"DlqKik*GE*～`u4H+LFJ(K-=%$%$</td><td>加载PFX格式的证书</td></tr>
<tr><td>**certificate load pem-chain** *cert-filename* **key-pair** ｛**dsa** | **rsa**｝**key-file** *key-filename***auth-code cipher** *auth-code*
例如：[HUAWEI-ssl-policy-https_der] **certificate load pem-chain chain-servercert.pem key-pair rsa key-file** chain-servercertkey.pem **auth-code cipher** 123456</td><td>加载PEM格式的证书链</td></tr>
</table>

3．配置HTTPS服务功能

HTTPS服务功能的配置与3.6.6节表3-24中第3点介绍的HTTP功能配置差不多，区别就是两点，这里不需要表3-24中第4步使能HTTP功能，另外在表3-24中第5步的 HTTP服务端口，这里要通过 http secure-server port port-number 命令配置的是HTTPS服务器监听端口，取值范围为443、1025～55535号端口。配置HTTPS服务器监听端口可以有效防止攻击者对HTTPS服务标准端口的访问，增加交换机的安全性。缺省情况下，安全HTTP服务器端监听端口号是443。其他配置完全一样，参见表3-24即可。

4．用户通过HTTPS登录交换机

完成以上配置后即可在PC上打开Web浏览器，在地址栏中输入“https://IP address”， 按回车键后将显示如图3-19所示的登录对话框。输入之前设置的Web网管账号和密码，输入验证码，并选择Web网管系统的语言，然后单击“登录”按钮或直接按回车键即可进入Web网管系统主页面。登录到Web网管后，可以对交换机进行管理和维护。

5．HTTPS Web网管登录管理

登录成功后可在交换机VRP命令行执行display ssl policy [policy-name] 命令查看配置的SSL策略及加载的数字证书；执行display http user命令查看当前在线用户信息；执行display http server命令，查看当前HTTPS服务器信息。

3.6.9 通过HTTPS Web网管登录交换机的配置示例

本示例拓扑结构如图 3-21 所示，现要求在作为 HTTP 服务器的交换机上部署 SSL策略，加载数字证书并使能HTTPS服务器功能后，用户可通过HTTPS登录到交换机，利用Web页面安全管理远程交换机。

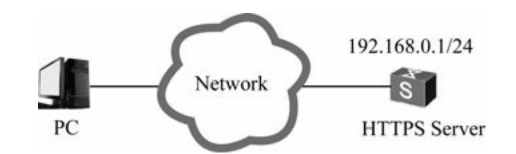


图3-21 通过HTTPS Web网管登录交换机的配置示例拓扑结构

1．基本配置思路

根据3.6.8节介绍的配置任务，结合本示例的具体要求，可得出本示例的基本配置思路如下。

（1）上传数字证书和Web网页文件：将PC上存储的数字证书、Web网页文件上传到作为HTTPS服务器的交换机上。

（2）加载数字证书和Web网页文件：将交换机存储器根目录下的数字证书文件复制到security子目录中，再配置SSL策略并加载数字证书和Web网页文件。

（3）使能HTTPS服务器功能，配置HTTPS服务属性参数。当然，也可不配置HTTPS服务属性参数，因为它们都有缺省值。

（4）在 AAA 视图下创建用于 HTTPS 登录的本地用户账户，并配置用户级别和对HTTP服务的支持。

（5）通过浏览器实现安全登录交换机。

2．具体配置步骤

（1）上传数字证书文件和 Web 网页文件。在此仅以通过 FTP 方式上传文件为例进行介绍。

<HUAWEI>system-view

[HUAWEI] sysname HTTPS-Server

[HTTPS-Server] ftp server enable #---使能FTP服务器功能

#---以下为配置FTP用户的验证信息、授权方式和授权目录，以便用户能通过FTP方式上传数字证书和Web网页文件。

[HTTPS-Server] aaa

[HTTPS-Server-aaa] local-userhuawei password cipher hello@123

[HTTPS-Server-aaa] local-userhuawei service-type ftp

[HTTPS-Server-aaa] local-userhuawei privilege level 15

[HTTPS-Server-aaa] local-userhuawei ftp-directory flash:

[HTTPS-Server-aaa] quit

[HTTPS-Server] quit

在用户终端PC的命令行提示符中执行ftp 192.168.0.1（192.168.0.1为交换机的管理IP地址）命令成功与交换机建立FTP连接后，然后使用put local-filename [remote-filename] 命令分别向交换机上传数字证书文件（包括服务器数字证书和服务器密钥这两个文件）和Web 网页文件。上传成功后，数字证书和Web 网页文件是保存在交换机存储器根目录下的。在交换机命令行下执行dir命令可看到成功上传的数字证书和Web网页文件（如粗体字部分所示）。

<HTTPS-Server>dir

Directory of flash:/

Idx Attr Size(Byte) Date Time FileName

0 -rw- 524,558 Apr 14 2011 16:24:39 private-data.txt

1 -rw- 1,302 Apr 14 2011 19:22:30 1_servercert_pem_rsa.pem

2 -rw- 951 Apr 14 2011 19:22:35 1_serverkey_pem_rsa.pem

3　drw-　-　Apr 09 2011 19:46:14　src

4　-rw-　421　Apr 09 2011 19:46:14　vrpcfg.zip

5　-rw-　1,308,478　Apr 14 2011 19:22:45　web001.7z

6　drw-　-　Apr 10 2011 01:35:54　logfile

7　-rw-　4　Apr 14 2011 04:56:35　snmpnotilog.txt

8　drw-　-　Apr 11 2011 16:18:53　security

9　drw-　-　Apr 13 2011 11:37:40　lam

...

65,233 KB total (7,289 KB free)

（2）配置SSL策略并加载数字证书。

#---以下为创建security目录，并将存储器根目录下的SSL数字证书文件复制到security目录中。

<HTTPS-Server>mkdir security/

<HTTPS-Server> copy 1_servercert_pem_rsa.pem security/

<HTTPS-Server> copy 1_serverkey_pem_rsa.pem security/

完成后在security目录下执行dir命令可看到复制成功的数字证书（如粗体字部分所示）。

<HTTPS-Server>cd security/

<HTTPS-Server>dir

Directory of flash:/security/

Idx　Attr　Size(Byte)　Date　Time　FileName

0　-rw-　1,302　Apr 13 2011 14:29:31　1_servercert_pem_rsa.pem

1　-rw-　951　Apr 13 2011 14:29:49　1_serverkey_pem_rsa.pem

65,233 KB total (7,287 KB free)

下面再来创建HTTPS服务器SSL策略，并通过 certificate load pem-cert命令加载PEM格式的数字证书（因为本示例中使用的是PEM格式的数字证书），包括服务器证书和服务器密钥这两个文件。

<HTTPS-Server>system-view

[HTTPS-Server] ssl policy http_server

[HTTPS-Server-ssl-policy-http_server] certificate load pem-cert 1_servercert_pem_rsa.pem key-pair rsa key-file 1_serverkey_pem_rsa.pem auth-code cipher123456

[HTTPS-Server-ssl-policy-http_server] quit

上述步骤成功配置后，在交换机命令行下执行display ssl policy命令可以看到加载 的数字证书详细信息，如下所示。

[HTTPS-Server] display ssl policy

SSL Policy Name: http_server

Policy Applicants:

Key-pair Type: RSA

Certificate File Type: PEM

Certificate Type: certificate

Certificate Filename: 1_servercert_pem_rsa.pem

Key-file Filename: 1_serverkey_pem_rsa.pem

Auth-code: 123456

MAC:

CRL File:

Trusted-CA File:

（3）加载新上传的Web 网页文件（如果使用交换机自带的Web 网页文件，则略过本步）。

[HTTPS-Server] http server loadweb001.7z

（4）使能HTTPS服务器功能，并创建HTTP用户，配置用户级别和对HTTP服务的支持。

[HTTPS-Server] http secure-server ssl-policyhttp_server

[HTTPS-Server] http secure-server enable

[HTTPS-Server] aaa

[HTTPS-Server-aaa] local-user admin password cipherhuawei

[HTTPS-Server-aaa] local-user admin privilege level 15

[HTTPS-Server-aaa] local-user admin service-typehttp

[HTTPS-Server-aaa] quit

（5）完成以上配置后就可以正式在PC终端通过HTTPS Web登录到交换机了。

在PC浏览器的地址栏中输入“https://192.168.0.1”，将显示登录对话框，如图3-19所示。然后正确输入HTTP用户名、密码和验证码，单击“登录”按钮或直接按回车键即可进入交换机的 Web 网管系统主页面。此时可在交换机的命令行界面下执行 display http server命令看到SSL策略名称和HTTPS服务器的状态（如粗体字部分显示）。

[HTTPS-Server] display http server

HTTP Server Status : disabled

HTTP Server Port : 80(80)

HTTP Timeout Interval : 20

Current Online Users : 1

Maximum Users Allowed : 5

HTTP Secure-server Status :enabled

HTTP Secure-server Port : 443(443)

HTTP SSL Policy :http_server

3.6.10 登录后的常用管理操作

用户通过Console口、Telnet或STelnet方式成功登录交换机后，在VRP系统命令行中用户除了可以对交换机进行业务配置外，还可以对当前登录用户以及交换机的基本功能执行如下管理任务。这些管理任务在日常的交换机登录用户管理中经常用到。

（1）显示在线用户。

（2）清除在线用户。

（3）设置切换用户级别的密码。

（4）切换用户级别。

（5）锁定用户配置权限。

（6）发送消息给其他用户界面。

（7）自动匹配上一级视图。

（8）锁定用户界面。

（9）允许在系统视图下执行用户视图命令。

（10）设置交换机允许的明文密码最小长度。

（11）配置告警级别。

下面分别予以介绍。

1．显示在线用户

用户登录系统后，可以使用display users [all] 查看每个用户界面的用户登录信息。该命令已在本章前面有介绍，不再赘述。

2．清除在线用户

当用户需要将某个登录用户与交换机的连接断开时，可以先使用display users命令来查看当前交换机上的用户登录信息，然后执行kill user-interface {ui-number | ui-type ui-number1 }命令清除指定用户界面下的在线用户（VRP 系统下的登录用户都是与用户界面一一对应的，所以清除用户时只需要断开对应的用户界面的连接，则相应的用户连接也就断开了，不是直接针对用户账户进行操作的）。在清除用户时系统会给出确认的。命令中的参数说明如下。

（1）ui-number：二选一参数，指定要清除用户的用户界面绝对编号，最小值为 0，最大值比系统支持的用户界面总数小1。不同交换机用户界面取值范围不同。

（2）ui-type ui-number1：二选一参数，指定要清除用户的用户界面类型和用户界面相对编号。

【示例 1】断开与 user-interface 0的连接。

<HUAWEI>kill user-interface 0

Warning: User interface con0 will be killd. Continue? [Y/N]y

3．设置用户级别的切换密码

如果当前用户级别较低，但是需要对高于用户级别的命令进行操作，用户可以执行super password [leveluser-level] [cipher password] 系统视图命令设置切换低级别用户到高级别用户的密码。输入的密码可以是明文或者密文，当不指定cipher password可选参数时，将采用交互方式输入明文密码；当指定cipher password可选参数时，既可以输入明文密码也可以输入密文密码，但都将以密文形式保存在配置文件中。明文密码为6～16个字符，区分大小写，密文密码为32个字符。输入的明文密码至少包含以下两种类型：大写字母、小写字母、数字及特殊字符（特殊字符不包括“？”和空格）。采用交互方式输入的密码不会在终端屏幕上显示出来。

【示例 2】设置从低级别登录的用户切换到level-3级别的切换密码为“abcdefg”。

<HUAWEI> system-view

[HUAWEI] super password level 3 cipher abcdefg

4．切换用户级别

用户由低级别切换到高级别时，可执行 super [level] 命令，在系统提示符下需要输入上面介绍的 super password命令设置好的切换密码。

如果输入的密码正确，将切换到更高级别；如果连续3次输入错误的口令，将退回用户视图，仍保持现有登录级别。当以低级别登录的用户通过super命令切换到高级别时，系统会自动发送trap信息，并记录在日志中；如果切换到的级别低于当前级别，则仅记录日志。

【示例 3】切换到用户级别3。在“Password”提示符下输入切换到对应级别的密码，但输入的密码不会在屏幕上显示的。

<HUAWEI> super 3

Password:

Now user privilege is 3 level, and only those commands whose level is equal to or less than this level can be used.

Privilege note: 0-VISIT, 1-MONITOR, 2-SYSTEM, 3-MANAGE

5．锁定用户配置权限

在多个用户同时登录系统进行配置时，有可能会出现配置冲突的情况。为了避免业务出现异常，可以配置权限互斥功能，保证同一时间只有一个用户可以配置。此时可执行 configuration exclusive命令锁定配置权限给当前操作用户（此命令可在所有视图下执行 ）。锁定用户配置权限后，可以明确地获取独享的配置权限，其他用户无法再获取到配置权限。

在锁定期间，执行锁定配置的用户可以进行查询、配置等操作，其他用户只能进行查询操作。也可在系统视图下执行configuration-occupied timeout timeout-value命令设置自行解锁时间间隔，取值范围为 1～7200s。缺省情况下，锁定间隔仅为30s。还可以执行 display configuration-occupied user命令查看当前锁定配置的用户信息。

【示例 4】查看当前锁定配置的用户信息。从输出信息可以看出当前锁定的用户值为34，使用的用户界面为VTY 0，锁定时间为 2013年 5月1日22时31分 36秒，以及其他信息，输出信息中的具体字段说明如表3-25所示。

<HUAWEI> display configuration-occupied user

User Index: 34

User Session Name: VTY0

User Name:

IP Address: 10.1.1.1

Locked Time: 22:31:36 05-01-2013

Last Configuration Time: 22:31:36 05-01-2013

The time out value of configuration right locked is: 30 second(s)

表3-25 display configuration-occupied user命令输出信息字段说明

字段	说明
User Index	显示被锁定的用户索引
User Session Name	显示被锁定的用户的会话名，VTY0～VTY14
User Name	显示被锁定用户的登录用户名称
IP Address	显示被锁定用户的用户 IP 地址，只对 VTY 连接用户有效
Locked Time	显示被锁定用户的锁定配置集的时间
Last Configuration Time	显示被锁定用户最后一次下发配置命令的时间
The time out value of configuration right locked is	显示配置权限锁定的时间

6．发送消息给其他用户界面

用户可以在当前的用户界面执行 send { all | ui-typeui-number | ui-number1 }命令发送消息给其他用户界面的用户，实现用户界面间的消息传递。命令中的参数和选项说明如下。

（1）all：多选一选项，向所有用户界面发送消息（包括当前没有登录的用户界面）。

（2）ui-type ui-number：多选一参数，向指定的相对编号用户界面发送消息。

（3）ui-number1：多选一参数，向指定的绝对编号用户界面发送消息。

执行本命令后，根据系统提示输入要传递的消息。使用Ctrl+Z组合键或回车键结束输入，使用Ctrl+C组合键中止本次操作。然后根据系统提示选择是否需要发送消息。按下Y键发送消息，按下N键取消发送。

【示例 5】向用户界面VTY 0发送消息。根据提示按下Y键确认后发送消息。发送后，通过VTY 0登录到交换机的用户就会收到这条信息。

<HUAWEI> send vty 0

Enter message, end with CTRL+Z or Enter; abort with CTRL+C:

Hello, good morning!

Warning: Send the message? [Y/N]:y

7．undo自动匹配上一级视图

可在系统视图下执行matched upper-view命令，以允许undo命令到上一级视图执行，但只对当前登录用户有效。这样，当用户在某个视图下执行非本视图注册的 undo命令时，系统将自动跳转到上一级视图搜索该undo命令。如果搜索成功，则该undo命令生效。当上级视图没有该 undo 命令时，系统就会自动向更上一级视图进行搜索，一直搜索到系统视图，不再向上搜索。

缺省情况下，undo命令不自动到上一级视图执行。但是要注意，非确实必要，不推荐配置这一功能，否则很容易出现配置错误，引起整个配置混乱。

8．锁定用户界面

当用户需要暂时离开操作终端时，为防止未授权的用户操作该终端界面，可以执行lock命令锁定当前用户终端界面，就像Windows系统中的用户桌面锁定一样。执行本命令后，根据系统提示输入锁定的密码，并确认密码（不会显示）。密码为6～16个字符，区分大小写。输入的密码至少包含以下几种类型：大写字母、小写字母、数字及特殊字符（特殊字符不包括“？”和空格）。

系统锁定后，如果想再次进入系统，必须先按回车键，然后根据提示输入锁定密码，只有在正确输入锁定密码后才可以解除锁定重新进入VRP系统。

【示例 6】锁定当前用户界面。

<HUAWEI>lock

Enter Password:

Confirm Password:

Info: The terminal is locked.

9．允许在系统视图下执行用户视图命令

缺省情况下，系统不允许在系统视图下执行用户视图命令，必须退出到用户视图才能成功执行。为了便于在不用切换视图的情况下执行用户视图下的命令，可通过在系统视图下执行 run command-line 命令配置系统视图下可执行的用户视图命 令。参数command-line为指定可在系统视图下执行的具体用户视图命令（不是命令格式本身）。

【示例 7】在系统视图下查看交换机上所有.cfg文件。

[HUAWEI] run dir*.cfg

Directory of cfcard:/

Idx Attr Size(Byte) Date Time FileName

0 -rw- 11,970 Mar 14 2012 19:11:22 9300_31.cfg

1 -rw- 12,033 Apr 22 2012 17:10:30 9300_31_new.cfg

509,256 KB total (118,784 KB free)

10．设置交换机允许的明文密码最小长度

就像在Windows服务器系统中，为了增加服务器系统的安全性，可在其密码策略中限制为用户账户配置的密码的最小密码长度一样，在华为VRP系统中也可以限制为用户配置的最小密码长度。可在系统视图

下执行 set password min-length length命令设置交换机允许的明文密码最小长度，使密码复杂性增加，从而提高交换机的安全性。参数length用来指定交换机允许的最小明文密码长度，取值范围为6～16的整数。缺省情况下，允许的最小明文密码长度是 6，可用undo set password min-length命令恢复交换机允许的最小明文密码长度为缺省的6 位。但该命令仅在 **S5700/7700/9700** 系列交换机上支持（**S9300**系列目前也不支持）。

说明

以上最小密码长度的设置仅对 local-user password cipher、set authentication password、super password和lock命令配置的明文密码具有最小长度限制作用，其他命令不受限制，例如OSPF、ISIS、RIP等协议的密码配置。执行本命令后，交换机上有关密码的明文配置必须符合此长度的限制。但对于恢复阶段配置的密码和已经生效的密码配置不会应用此长度限制 。

3.6.11 常见配置错误分析与排除

本节要介绍Telnet和STelnet登录失败的故障分析与排除方法。

1．Telnet登录失败的故障分析与排除

出现 Telnet 登录失败的原因可能有很多，可按照以下流程进行分析与排除（从Console口登录到交换机执行以下命令）。

（1）查看所使用的VTY用户界面视图是否允许支持Telnet服务。

执行user-interface vty命令进入对应用户界面视图，然后执行display this命令查看对应VTY用户界面的protocol inbound命令配置项是否为 telnet或者 all（缺省情况下为 all）。如果不是，执行protocol inbound { telnet | all }命令修改配置，以允许 telnet类型用户接入交换机。

（2）查看登录交换机的用户数是否到达了上限。

执行 display users命令查看当前的VTY通道是否全部被占用。缺省情况下，VTY通道允许的最大用户数是 5个，可执行display user-interface maximum-vty命令查看当前 VTY 通道允许的最大用户数。如果当前的用户数已经达到上限，可以执行命令user-interface maximum-vty 15，将VTY通道允许的最大用户数扩展到15个。

（3）查看交换机上VTY用户界面视图下是否正确配置了ACL。

执行user-interface vty命令进入对应用户界面视图，然后执行display this命令查看对应的VTY用户界面是否配置了ACL限制，如果配置了ACL限制，请记录该ACL编号。然后执行display aclacl-number命令查看该访问控制列表中是否deny了Telnet客户端的地址。如果是，则在ACL视图下执行undo rule rule-id命令删除该deny规则，再执行 rulepermit source source-ip-address soucer-wildcard命令修改访问控制列表，以允许客户端的IP地址访问。

（4）查看VTY用户界面视图下是否正确设置登录验证。

如果使用authentication-mode password命令配置了VTY线路下的登录验证方式为密码验证方式，则必须在登录时正确输入此密码；如果使用authentication-mode aaa命令设置验证方式为AAA验证，则必须使用local-user user-name password命令创建AAA本地用户，并配置密码。

2．STelnet登录失败

如果出现STelnet登录到SSH服务器失败，则需要按照以下流程进行故障分析与排除（从Console口登录或者Telnet方式登录到交换机执行以下命令）。

（1）查看VTY用户界面视图下是否允许支持SSH服务。

执行user-interface vty命令进入对应用户界面视图，然后执行display this命令查看VTY用户界面的protocol inbound是否为ssh或者all。如果不是，则执行protocol inbound { ssh | all }命令修改配置，以允许

STelnet类型用户接入交换机。

（2）查看登录SSH服务器端的用户数是否到达了上限。

执行 display users命令查看当前的VTY通道是否全部被占用。缺省情况下，VTY通道允许的最大用户数是 5个，可以先执行display user-interface maximum-vty命令查看当前 VTY 通道允许的最大用户数。如果当前的用户数已经达到上限，可以执行命令user-interface maximum-vty 15，将VTY通道允许的最大用户数扩展到 15个。

（3）查看SSH服务器端上VTY用户界面下是否绑定了ACL。

执行 user-interface vty 命令进入对应的 SSH 用户会使用的界面视图，然后执行display this命令查看VTY用户界面是否配置了ACL限制，如果配置了ACL限制，请记录该ACL编号。再执行display aclacl-number命令查看该访问控制列表中是否deny了Telnet客户端的地址。如果是，则在ACL视图下执行undo rule rule-id，命令删除该deny规则，再执行 rulepermit source source-ip-address soucer-wildcard命令修改访问控制列表，以允许客户端的IP地址访问。

（4）查看SSH客户端和服务器上SSH版本是否兼容。

执行display ssh server status命令查看SSH版本信息。如果使用的是SSHv1版本的客户端登录服务器，则需要执行 ssh server compatible-ssh1x enable命令配置SSH服务器兼容SSHv1版本。

（5）查看SSH服务器端的SSH服务是否启动。

执行 display ssh server status命令查看 SSH服务器端配置信息。如果 SSH服务器功能没有使能，则执行如下 stelnet server enable命令使能 SSH服务器端的 STelnet服务。

（6）查看在SSH服务器端是否配置了RSA或DSA公钥。

当交换机作为 SSH服务器时必须配置本地密钥对，执行 display rsa local-key-pair public或display dsa local-key-pair public命令查看当前服务器端密钥对信息。如果显示信息为空，则表明没有配置服务器端密钥对，执行 rsa local-key-pair create 或 dsa local-key-pair create命令创建。

（7）查看SSH服务器端上是否配置了SSH用户。

执行display ssh user-information命令查看SSH用户的配置信息。如果不存在配置信息，请在系统视图下执行 ssh user、ssh user authentication-type和 ssh user service-type命令新建SSH用户，并正确配置SSH用户的验证方式和SSH用户的服务方式。

（8）查看SSH客户端是否使能了首次验证功能。

在系统视图下执行 display this命令查看 SSH客户端是否使能了 SSH客户端首次验证功能。如果没有使能，则STelnet客户端第一次登录SSH服务器时由于对SSH服务器的 RSA公钥有效性检查失败，而导致登录服务器失败，此时需要执行 ssh client first-time enable命令使能 SSH客户端首次验证功能。使能了 SSH客户端首次认证功能后，当STelnet/SFTP客户端第一次登录SSH服务器时不对SSH服务器的RSA或DSA公钥进行有效性检查，因为此时 STelnet/SFTP 客户端还没有保存 SSH 服务器的 RSA或DSA公钥。

3.7 远程文件管理

交换机中所有的文件保存在存储器中，可通过多种方式实现对存储器中本地文件的管理。前面介绍的通过Console口、MiniUSB口或者Telnet、STelnet方式直接登录到交换机的VRP系统后，就可以对交换机上的VRP文件系统进行全面的管理，具体的管理方法参见本书第2章2.4节。除此之外，还可通过一些文件传输协议连接、访问交换机，进行一些基本的文件管理工作和文件传输操作。同时也可以将当前交换机作为客户端，通过多种方式实现对其他交换机文件的访问。因篇幅原因，在此仅介绍将交换机作为服务器端进行的

远程文件管理和传输配置。

3.7.1 文件管理方式的支持

目前，在华为S系列交换机中用户可以通过Console口或MiniUSB口、Telnet、STelnet登录方式直接登录系统，通过FTP、TFTP、SFTP、SCP或FTPS方式进行远程文件管理。交换机在进行文件管理的过程中，可以分别充当服务器和客户端的角色。

（1）交换机作为服务器：可以从终端访问交换机，实现对本交换机文件的管理，以及与终端间的文件传输操作。

（2）交换机作为客户端访问其他交换机（服务器）：可以实现管理其他交换机上的文件，以及与其他交换机间进行文件传输操作。

对于TFTP方式，交换机只支持客户端功能；对于FTP、SFTP、SCP以及FTPS方式，交换机均支持服务器与客户端功能。以上这些文件管理方式的应用场景，优缺点如表3-26所示，用户可以根据需求选择其中一种。

表3-26 文件管理方式比较

文件管理方式	应用场景	优点	缺点
直接登录系统	通过 Console 口、Telnet 或 STelnet 方式登录交换机，对存储器、目录和文件进行管理。特别是对存储器的操作需要通过此种方式	对存储器、目录和文件的管理直接通过登录交换机完成，方便快捷	只是对本交换机进行文件操作，**无法进行文件的传输**
FTP	适用于对网络安全性要求不是很高的文件传输场景中，广泛用于版本升级等业务中	配置较简单，**支持文件传输以及文件、目录的操作**；可在两个不同文件系统主机之间传输文件。具有授权和验证功能	明文传输数据，存在安全隐患
TFTP	在网络条件良好的实验室局域网中，可以使用 TFTP 进行版本的在线加载和升级。适用于客户端和服务器之间不需要复杂交互的环境	所占的内存要比 FTP 小，只支持文件传输	交换机**只支持 TFTP 客户端功能；只支持文件传输，不支持交互操作**，TFTP 没有授权和验证，且是明文传输数据，存在安全隐患，易于网络病毒传输以及被黑客攻击
SFTP	适用于网络安全性要求高的场景，目前被广泛用于日志下载、配置文件备份等业务中	数据进行了严格加密和完整性保护，安全性高。**支持文件传输及文件、目录的操作**	配置较复杂。在交换机上可以同时配置 SFTP 功能和普通 FTP 功能。（这一点与 FTPS 方式相比：FTPS 是不可以同时提供 FTPS 和普通 FTP 功能的）
SCP	适用于网络安全性要求高，且文件上传下载效率高的场景	在安全性方面，与 SFTP 一样。**在客户端与服务器连接的同时完成文件的上传、下载操作**（即连接和复制操作使用一条命令完成），效率较高	配置较复杂（与 SFTP 方式的配置非常类似），**但不支持交互操作**
FTPS	适用于网络安全性要求高，且不提供普通 FTP 功能的场景	利用数据加密、身份验证和消息完整性验证机制，为基于 TCP 可靠连接的应用层协议提供安全性保证	配置较复杂，需要预先从 CA 处获得一套证书。如果配置了 FTPS 服务，则需关闭普通 FTP 服务功能

说明

因为像通过Console口、MiniUSB口、Telnet、STelnet方式登录交换机的方法在 本章前面已有全面介绍，再加上登录后的本地文件系统管理方法已在第2章详细介绍，故在此不再赘述，参见即可。本节仅介绍通过PC终端的FTP、SFTP、SCP、FTPS客户端软件访问交换机（此时交换机作为服务器端）进行远程文件

管理的配置方法。

3.7.2 通过**FTP**进行文件操作

用户可以使用FTP协议在本地与远程终端之间进行文件操作，在版本升级等文件业务操作中此协议广泛应用。配置前需要确保终端与交换机之间路由可达和终端支持FTP客户 端软件。但使用FTP协议存在安全风险，建议使用SFTP或FTPS方式进行文件操作。

1．配置任务

通过FTP进行文件操作的配置任务如下所示（第1～3步之间没有严格的配置顺序）：

（1）配置FTP服务器功能及参数：使能FTP服务器，配置FTP服务器属性参数，如端口号、源IP地址、超时断连时间。

（2）配置FTP本地用户：配置本地用户的服务类型、用户级别及授权访问目录等。

（3）（可选）配置FTP访问控制：配置用于控制FTP用户访问的ACL列表，提高FTP访问的安全性。仅在需要通过**ACL**进行**FTP**访问控制时选用 。

（4）用户通过FTP访问交换机：从终端通过FTP访问交换机。

与FTP文件操作方式的相关参数缺省配置为：FTP服务器功能关闭，监听21号TCP端口，无FTP本地用户。

下面介绍具体的配置任务。

2．配置FTP服务器功能及参数

FTP服务器功能的使能和参数配置比较简单，具体如表3-27所示。

表3-27 FTP服务器功能使能及参数的配置步骤

步骤	命令	说明
1	**system-view** 例如：<HUAWEI> **system-view**	进入系统视图
2	**ftp server port** *port-number* 例如：[HUAWEI] **ftp server port** 1088	（可选）指定 FTP 服务器端口号，取值范围为 21 或 1 025～55 535 的整数。缺省情况下，FTP 服务器端监听端口号是 21，可用 **undo ftp server port** 命令恢复缺省值 【说明】当服务器正在监听的端口号是 21 时，FTP 客户端登录时可以不指定端口号，因为 21 号端口是 FTP 服务器的缺省端口；如果是其他监听端口号，FTP 客户端登录时必须指定对应的端口号。但客户端的端口号必须与服务器端指定的端口号一致。但在变更端口前需要确保 FTP 服务处于非使能状态，否则需要先执行 **undo ftp server** 命令关闭服务。使用本命令变更端口号后需要执行 **ftp server enable** 命令重新使能 FTP 服务
3	**ftp server enable** 例如：[HUAWEI]**ftp server enable**	在交换机上使能 FTP 服务器功能。缺省情况下，交换机上的 FTP 服务器功能是关闭的，可用 **undo ftp server** 命令关闭交换机的 FTP 服务器功能。关闭 FTP 服务器功能后，未登录的用户将无法登录 FTP 服务器。已经登录到该 FTP 服务器上的用户，除了退出登录的操作外，不能再执行任何操作
4	**ftp server-source** { **-a** *source-ip-address* | **-i** *interface-type interface-num* } 例 如：[HUAWEI]**ftp server-source-i** loopback0	（可选）指定 FTP 服务器的源地址或源接口，实现对交换机进出报文的过滤，保证安全性。命令中的参数说明如下： • *source-ip-address*：二选一参数，用来指定 FTP 服务器源 IP 地址 • *interface-type interface-num*：二选一参数，用来指定 FTP 服务器的源接口 但 FTP 服务器端指定的源地址**只能是交换机的 LoopBack 接口 IP 地址或 LoopBack 接口**。配置了服务器的源地址后，登录服务器时所输入的服务器地址必须与该命令中配置的一致，否则无法成功登录。**如果在配置此命令前，FTP 服务已经使能，在则在配置本命令后 FTP 服务将重新启动** 缺省情况下，FTP 服务器发送报文的源地址为 0.0.0.0（代表任意 IP 地址），可用 **undo ftp server-source** 命令恢复 FTP 服务器发送报文的源地址为缺省值

（续表）

步骤	命令	说明
5	**ftp timeout** *minutes* 例如：[HUAWEI] **ftp timeout** 20	（可选）配置 FTP 连接最大空闲等待时间，取值范围为（1～35791）整数分钟 【说明】用户登录到 FTP 服务器后如果连接异常中断或用户非正常中断连接，FTP 服务器是无法知道的，因而连接仍保持着。为防止这类情况发生，使用连接空闲时间，当连接在一定时间内没有进行命令交互，FTP 服务器即可认为连接已经失效，而断开连接 缺省情况下，连接空闲时间为 30min，可用 **undo ftp timeout** 命令恢复缺省的连接空闲时间

【示例 1】设置FTP服务器的源地址为LoopBack0接口。在配置应用前系统会先给出一个FTP服务器将重启的警告提示。确认后才正式应用配置。

<HUAWEI>system-view

[HUAWEI] ftp server-source -i loopback0

Warning: To make the server source configuration take effect, the FTP server will be restarted. Continue? [Y/N]: y

Info: Succeeded in setting the source interface of the FTP server to LoopBack0.

Info: Succeeded in starting the FTP server.

3．配置FTP本地用户

当用户通过FTP进行文件操作时，需要在作为FTP服务器的交换机上配置本地用户名及口令（进行的是AAA验证方式）、指定用户的服务类型以及可以访问的目录，否则用户将无法通过FTP访问交换机。具体的配置步骤如表3-28所示。

表3-28 FTP本地用户的配置步骤

步骤	命令	说明
1	**system-view** 例如：<HUAWEI> **system- view**	进入系统视图
2	**aaa** 例如：[HUAWEI] **aaa**	进入 AAA 视图
3	**local-user** *user-name* **password cipher** *password* 例如：[HUAWEI-aaa] **local-user** winda **password cipher** 123456	配置本地用户名和密码。缺省情况下，系统中没有本地用户，**也不支持 FTP 匿名访问**。本命令在本章前面已多次介绍，参见即可
4	**local-user** *user-name* **privilege level** *level* 例如：[HUAWEI-aaa] **local-user** winda **privilege level** 5	配置本地用户级别。本命令在本章前面已多次介绍，参见即可。但此处**必须将用户级别配置在 3 级或 3 级以上，否则 FTP 连接将无法成功**。 缺省情况下，本地用户（如 Telnet 用户、SSH 用户）的优先级由对应的用户界面优先级决定，可用 **undo local-user** *user-name* **privilege level** 命令将指定的本地用户的优先级恢复为缺省配置
5	**local-user** *user-name* **service-type ftp** 例如：[HUAWEI-aaa] **local-user** winda **service-type ftp**	配置本地用户的服务类型为 FTP。缺省情况下，本地用户可以使用所有的接入类型，包括 8021x（支持 802.1x 认证的用户）、bind（IP 会话用户）、ftp（FTP 连接用户）、http（HTTP 连接用户）、ppp（PPP 连接用户）、ssh（STelnet 连接用户）、telnet（Telnet 连接用户）、terminal（Console 口或者 MiniUSB 口连接用户）和 web（Web 认证用户），可用 **undo local-user** *user-name* **service-type** 命令将指定的本地用户的接入类型恢复为支持所有接入类型

（续表）

步骤	命令	说明
6	**local-user** *user-name* **ftp-directory** *directory* 例如：[HUAWEI-aaa] **local-user** winda **ftp-directory** flash:/	配置本地用户的 FTP 授权访问目录（包括完整的目录路径），为1～64 个字符，不支持空格，**区分大小写** 当有多个 FTP 用户且有相同的授权目录时，可以执行 **set default ftp-directory** *directory* 命令为 FTP 用户配置缺省工作目录。此时，不需要通过本命令为每个用户配置授权目录 缺省情况下，本地用户的 FTP 目录为空，可用 **undo local-user** *user-name* **ftp-directory** 命令将指定的本地用户的 FTP 目录删除

【示例 2】设置本地用户hello@huawei.net的优先级为6。

<HUAWEI>system-view

[HUAWEI] aaa

[HUAWEI-aaa] local-user hello@huawei.net privilege level 6

【示例 3】设置本地用户hello@huawei.net的FTP目录为flash:/。

<HUAWEI>system-view

[HUAWEI] aaa

[HUAWEI-aaa] local-userhello@huawei.net ftp-directory flash:/

4．（可选）配置FTP访问控制

用户可以配置FTP访问控制列表，实现只允许指定的客户端登录到交换机，以提高安全性。当ACL中的规则选择permit选项时，则允许指定源IP地址的其他交换机与本交换机建立FTP连接；当ACL中的规则选择deny选项时，则拒绝其他交换机与本交换机建立FTP连接；当ACL未配置规则时，则允许任何其他交换机与本交换机建立FTP连接。具体配置步骤如表3-29所示。

表3-29 FTP访问控制的配置步骤

步骤	命令	说明
1	**system-view** 例如：<HUAWEI> **system-view**	进入系统视图
2	**acl** [**number**] *acl-number* 例如：[HUAWEI] **acl** 2001	进入 ACL 视图。命令中的参数和选项说明如下。 （1）**number**：可选项，指定是由数字标识的一个基本 ACL （2）*acl-number*：指定用于控制 FTP 用户访问的基本 ACL 的编号，取值范围为 2000～2999 的整数（规则中的源 IP 地址就是允许或者禁止进行 FTP 访问的用户计算机的 IP 地址或所在网段） 可由 **undo acl** { [**number**] *acl-number* | **all** }删除指定的 ACL
3	**rule** [*rule-id*] { **deny** | **permit** } [**source** {*source-address source-wildcard* | **any** } | **fragment**| **logging**| **time-range** *time-name*] 例如：[HUAWEI-acl-basic-2001] **rule permit source** 192.168.32.1 0	配置 ACL 规则。命令中的参数和选项说明如下。 （1）*rule-id*：指定 ACL 的规则 ID。 如果指定 ID 的规则已经存在，则会在旧规则的基础上叠加新定义的规则，相当于编辑一个已经存在的规则（**通过这种方法可以修改现有 ACL 规则**）；如果指定 ID 的规则不存在，则使用指定的 ID 创建一个新规则，并且按照 ID 的大小决定规则插入的位置。 如果不指定 ID，则增加一个新规则时自动会根据设置的 ID 步长为这个规则分配一个 ID，ID 按照大小排序，规则 ID 的步长由 **step** *step-value* 命令指定，缺省步长为 5 （2）**deny**：二选一选项，指定拒绝符合条件的数据包 （3）**permit** 二选一选项，指定允许符合条件的数据包

（续表）

步骤	命令	说明
3	**rule** [*rule-id*] { **deny** | **permit** } [**source** {*source-address source-wildcard* | **any** } | **fragment**| **logging**| **time-range** *time-name*] * 例如：[HUAWEI-acl-basic-2001] **rule permit source** 192.168.32.1 0	（4）*source-address source-wildcard*：二选一参数，指定数据包源 IP 地址和源 IP 地址通配符掩码。*source-address* 为点分十进制形式，或用 **any** 代表任意源地址 0.0.0.0；*source-wildcard* 为点分十进制形式，数值上是源地址掩码的反掩码形式。当目的地址是 **any** 时，通配符是 255.255.255.255；当目的地址是主机时，通配符是 0 （5）**any**：二选一选项，表示数据包的任意源地址 （6）**fragment**：可多选项，指定该规则是否仅对非首片分片报文有效。当包含此选项时表示该规则仅对非首片分片报文有效 （7）**logging**：可多选项，指定把 ACL 的匹配信息写进日志 （8）*time-name*：可多选参数，指定 ACL 规则生效的时间段。其中 *time-name* 表示 ACL 规则生效的时间段名称，长度范围为 1～32 个字符 可用 **undo rule** *rule-id* [**fragment** | **logging** | **source** | **time-range**] *命令删除一个基本 ACL 规则
4	**quit** 例如：[HUAWEI-acl-basic-2001] **quit**	退出基本 ACL 视图，返回系统视图
5	**ftp acl** *acl-number* 例如：[HUAWEI] **ftp acl** 2001	在 FTP 的交换机连接中应用指定的 ACL，设置允许哪些客户端访问本 FTP 服务器

【示例 4】在ACL 2001中增加一条规则，允许源地址为主机地址 192.168.32.1的报文通过。

<HUAWEI>system-view

[HUAWEI] acl 2001

[HUAWEI-acl-basic-2001] rule permit source192.168.32.1 0

5．用户通过FTP访问交换机

完成以上配置后，用户就可以从终端通过FTP协议访问交换机。此时用户可以选择使用Windows命令行提示符或第三方软件进行FTP访问操作。在此仅以Windows命令行提示符为例进行介绍。

方法很简单，仅需在Windows命令提示符下输入 ftp 192.168.150.208（假设交换机的IP地址为192.168.150.208）命令，通过FTP协议访问交换机。然后根据提示输入用户名和口令，按回车键，当出现FTP客户端视图的命令行提示符，如ftp>，此时用户进入了FTP服务器的工作目录，就可以进行各种基于FTP协议的文件管理，如上传、下载交换机系统软件和配置文件等。

C:\Documents and Settings\Administrator> ftp 192.168.150.208

Connected to 192.168.150.208.

220 FTP service ready.

User(192.168.150.208:(none)):huawei

331 Password required for huawei.

Password:

230 User logged in.

ftp>

6．通过FTP命令进行文件操作

用户成功访问担当FTP服务器的华为S系列交换机后，在PC终端的命令提示符下可以通过FTP命令进行文件操作，包括目录操作、文件操作、配置文件传输方式、上传或下载文件，查看FTP命令在线帮助等，如表3-30所示。有关文件和目录管理命令的使用方法参见本书第2章2.4节。但用户的操作权限受限于服务器上对该用户的用户优先级设置。

表3-30 通过FTP命令可进行的文件操作

<table>
<tr><th>命令</th><th colspan="2">说明</th></tr>
<tr><td>cd remote-directory</td><td colspan="2">改变服务器上的工作路径</td></tr>
<tr><td>cdup</td><td colspan="2">改变服务器的工作路径到上一级目录</td></tr>
<tr><td>pwd</td><td colspan="2">显示服务器当前的工作路径</td></tr>
<tr><td>lcd [local-directory]</td><td colspan="2">显示或者改变客户端的工作路径到指定目录。与 pwd 命令不同的是，lcd 命令执行后显示的是客户端的本地工作路径，而 pwd 显示的则是远端服务器的工作路径</td></tr>
<tr><td>mkdir remote-directory</td><td colspan="2">在服务器上创建指定目录。创建的目录可以为字母和数字等的组合，但不可以为<、>、?、\、: 等特殊字符</td></tr>
<tr><td>rmdir remote-directory</td><td colspan="2">在服务器上删除指定目录</td></tr>
<tr><td>dir/ls [remote-filename [local-filename]]</td><td colspan="2">显示服务器上指定目录或文件的信息。ls 命令只能显示出目录/文件的名称，而 dir 命令可以查看目录/文件的详细信息，如大小，创建日期等
如果指定远程文件时没有指定路径名称，那么系统将在用户的授权目录下搜索指定的文件</td></tr>
<tr><td>delete remote-filename</td><td colspan="2">删除服务器上指定文件</td></tr>
<tr><td>put local-filename [remote-filename]
或
mput local-filenames</td><td colspan="2">上传指定的单个或多个文件。put 命令是上传单个文件；mput 命令是上传多个文件</td></tr>
<tr><td>get remote-filename [local-filename]
或
mget remote-filenames</td><td colspan="2">下载指定的单个或多个文件。get 命令是下载单个文件；mget 命令是下载多个文件</td></tr>
<tr><td>ascii</td><td>配置传输文件的数据类型为 ASCII 模式</td><td rowspan="2">（二选一）缺省情况下，文件传输方式为 ASCII 模式
传输文本文件使用 ASCII 方式；传输程序、数据库文件等使用二进制模式</td></tr>
<tr><td>binary</td><td>配置传输文件的数据类型为二进制模式</td></tr>
<tr><td>passive</td><td>配置文件传输方式为被动方式</td><td rowspan="2">（二选一）缺省情况下，数据传输方式是被动方式。主动模式是从服务器端向客户端发起；被动模式是客户端向服务器端发起连接</td></tr>
<tr><td>undo passive</td><td>配置文件传输方式为主动方式</td></tr>
<tr><td>remotehelp [command]</td><td colspan="2">查看 FTP 命令的在线帮助</td></tr>
<tr><td>prompt</td><td colspan="2">使能系统的提示功能。缺省情况下，不使能信息提示</td></tr>
<tr><td>verbose</td><td colspan="2">打开 verbose 开关。如果打开 verbose 开关，将显示所有 FTP 响应，包括 FTP 协议信息，以及 FTP 服务器返回的详细信息</td></tr>
</table>

7．FTP访问管理

可以在不退出当前FTP客户端视图的情况下，通过user user-name [password] 命令 以其他的用户名登录到FTP服务器交换机上。所建立的FTP连接，与执行ftp命令建立的FTP连接完全相同。但更改当前的登录用户后，原用户与服务器的连接将断开。

如果要断开与FTP服务器的连接，用户可以在FTP客户端视图中选择不同的命令断开与FTP服务器的连接：通过bye或quit命令可以终止与服务器的连接，并退回到用户视图；通过close或disconnect命令可以终止与服务器的连接，并退回到FTP客户端视图。

还可使用display ftp-server任意视图命令查看FTP服务器的配置和状态信息；使用display ftp-users 任意视图命令查看登录的 FTP 用户信息；使用 display acl { acl-number | all }任意视图命令查看访问控制列表的配置信息。

3.7.3 通过FTP进行文件操作的配置示例

本示例拓扑结构如图3-22所示，PC与交换机之间的路由可达，10.136.23.5是交换机的管理口IP地址。现在交换机需要升级VRP系统，将交换机作为FTP服务器，从终端PC将VRP系统软件上传至交换机，且保存当前交换机的配置文件到终端进行备份（也就是执行文件下载操作）。

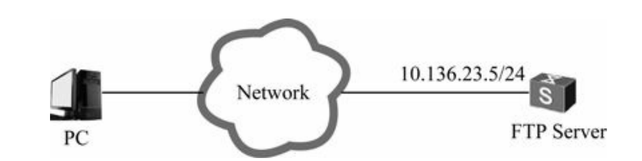


图3-22 通过FTP进行文件操作的配置示例拓扑结构

1．基本配置思路

根据3.7.2介绍的FTP文件操作配置任务，以及本示例的具体要求，可以得出本示例的基本配置思路如下。

（1）配置交换机的FTP服务器功能及FTP用户信息（包括用户名及密码、用户级别、服务类型、授权目录）。

（2）保存交换机当前配置文件，以便下载到终端备份。

（3）从安装了FTP客户端软件的终端PC上通过FTP协议连接担当FTP服务器的交换机。

（4）将要用于升级的 VRP 系统软件上传至交换机存储器根目录中，然后下载在交换机上保存的配置文件备份到PC终端。

2．具体配置步骤

（1）配置交换机的FTP服务器功能及FTP用户信息，包括用户名、密码、用户级别、FTP服务的支持和授权访问的目录。

```
<HUAWEI> system-view
[HUAWEI] ftp server enable
[HUAWEI] aaa
[HUAWEI-aaa] local-userhuawei password cipher huawei@123    #---创建用户huawei，并设置其密码为huawei@123
[HUAWEI-aaa] local-user huawei privilege level 15    #---设定用户huawei具有最高的15级权限
[HUAWEI-aaa] local-userhuawei service-type ftp    #---设定用户huawei支持的服务类型为 ftp
[HUAWEI-aaa] local-userhuawei ftp-directory flash:/    #---授权访问 flash:根目录
[HUAWEI-aaa] quit
[HUAWEI] quit
```

（2）保存交换机当前配置文件。

```
<HUAWEI> save
```

（3）从终端PC通过FTP协议连接交换机（交换机的管理口IP地址为10.136.23.5），输入用户名huawei和密码huawei@123。下面仅以Windows XP操作系统的提示符为例进行介绍。

```
C:\Documents and Settings\Administrator> ftp 10.136.23.5
Connected to 10.136.23.5.
220 FTP service ready.
User (10.136.23.5:(none)): huawei
331 Password required for huawei.
Password:
230 User logged in.
ftp>
```

（4）通过put命令将VRP系统软件（假设VRP系统软件名为devicesoft.cc）上传至交换机。

ftp>put devicesoft.cc

200 Port command okay.

150 Opening ASCII mode data connection for devicesoft.cc.

226 Transfer complete.

ftp:发送 23876556 字节，用时 25.35Seconds 560.79Kbytes/sec.

（5）使用get命令将交换机上的配置文件（假设配置文件名为vrpcfg.zip）下载到终端PC的当前目录（可以在PC机上保存至另外目录下）进行备份。

ftp>get vrpcfg.zip

200 Port command okay.

150 Opening ASCII mode data connection for vrpcfg.zip.

226 Transfer complete.

ftp:收到 1257字节，用时 0.03Seconds 40.55Kbytes/sec.

（6）检查配置结果。在交换机中执行 dir 命令，查看系统软件是否上传至交换机存储器的根目录下。如果上传成功，可以在输出信息中见到它（如粗体字部分）。

<HUAWEI> dir

Directory of flash:/

Idx Attr Size(Byte) Date Time FileName

0 -rw- 14 Mar 13 2012 14:13:38 back_time_a

1 drw- - Mar 11 2012 00:58:54 logfile

2 -rw- 4 Nov 17 2011 09:33:58 snmpnotilog.txt

3 -rw- 11,238 Mar 12 2012 21:15:56 private-data.txt

4 -rw- 1,257 Mar 12 2012 21:15:54 vrpcfg.zip

5 -rw- 14 Mar 13 2012 14:13:38 back_time_b

6 -rw- 23,876,556 Mar 13 2012 14:24:24 devicesoft.cc

7 drw- - Oct 31 2011 10:20:28 sysdrv

8 drw- - Feb 21 2012 17:16:36 compatible

9 drw- - Feb 09 2012 14:20:10 selftest

10 -rw- 19,174 Feb 20 2012 18:55:32 backup.cfg

11 -rw- 23,496 Dec 15 2011 20:59:36 20111215.zip

12 -rw- 588 Nov 04 2011 13:54:04 servercert.der

13 -rw- 320 Nov 04 2011 13:54:26 serverkey.der

14 drw- - Nov 04 2011 13:58:36 security

..

65,233 KB total (7,289 KB free)

至此，整个配置任务已全部完成。

3.7.4 通过SFTP进行文件操作

SFTP是SSH协议的一部分，需要通过**VTY**用户界面进行连接（而**FTP**协议不需要通过**VTY**用户界面连接）。SFTP使得用户终端可以在SSH协议的基础上与远端交换机进行安全连接，同时在远程系统升级、日志下载等场景下增加了数据传输的安全性。

在配置通过 SFTP 进行文件操作之前，也需要确保终端与交换机之间有可达路由，且终端上已安装SSH客户端软件（如putty或OpenSSH软件）。通过SFTP进行文件操作的配置任务如下（第1～3步之间没有严格的配置顺序）。

（1）配置SFTP服务器功能和参数：包括服务器本地密钥对生成、SFTP服务器功能的使能及服务器参数的配置：监听端口号、密钥对更新时间、SSH验证超时时间、SSH验证重试次数等。

因为SFTP与STelnet一样都是使用SSH服务，所以本步配置与本章前面3.6.4节的第2点介绍的STelnet服务器功能和参数配置差不多，唯一不同的是在表3-16中的第3步，这里要启用的是SFTP服务，使用的命令是sftp server enable，其他配置完全一样，参见表3-16即可。

（2）配置SSH用户登录的用户界面：包括VTY用户界面的用户验证方式、VTY用户界面支持SSH协议及其他基本属性。

本项配置任务的配置方法参见本章 3.5.1～3.5.4 节（在表 3-13 中第 8 步要通过protocol inbound ssh命令配置对应的VTY用户界面支持SSH服务），验证方式的配置参见本章3.5.5节，建议采用AAA验证方式（也可以采用其他验证方式 ）。

（3）配置SSH用户：包括SSH用户的创建、验证方式、服务方式、SFTP服务授权目录等。

本项配置任务的配置也与3.6.4节介绍的STelnet登录第3点中的SSH用户配置差不多，不同的是要在表3-17第4步中使用 ssh userusername service-type { sftp |all }命令配置SSH用户支持SFTP服务，缺省情况下，SSH用户的服务方式是空，即不支持任何服务方式。另外还要通过 ssh userusername sftp-directorydirectoryname命令配置SSH用户的SFTP服务授权目录，缺省情况下，SSH用户的SFTP服务授权目录是flash:。其他配置参见表3-16。

在SSH用户验证方面：

① 如果要对 SSH用户进行 password、password-dsa或 password-rsa验证，按 3-18进行配置。

② 如果要对SSH用户进行 dsa、rsa、password-dsa或 password-rsa验证，按表 3-19进行配置。

因为前面各项配置任务在本章前面都有相关介绍，且不同之处在以上各配置任务中均有说明，故不再赘述。下面仅介绍用户通过 SFTP 协议访问交换机，以及访问成功后可进行的文件操作和访问管理。

1．用户通过SFTP协议访问交换机

从终端通过SFTP访问交换机需要在终端上安装SSH客户端软件。此处以使用第三方软件OpenSSH和Windows命令行提示符为例进行配置。使用OpenSSH软件从终端访问交换机时需要使用OpenSSH的命令，命令的使用可以参见该软件的帮助文档。只有安装了OpenSSH软件后，Windows命令行提示符才能识别OpenSSH相关命令。

操作方法是先进入 Windows 的命令行提示符，然后在命令行中输入 sftp username命令（如 sftp sftpuser@10.136.23.5），按回车键后即开始与交换机进行连接，按照提示正确输入用户名和密码。连接成功后即出现sftp>提示符，表示用户已进入了SFTP服务器的工作目录。如下所示：

```
C:\Documents and Settings\Administrator> sftp sftpuser@10.136.23.5
Connecting to 10.136.23.5. .
The authenticity of host '10.136.23.5 (10.136.23.5)' can't be established.
RSA key fingerprint is 46:b2:8a:52:88:42:41:d4:af:8f:4a:41:d9:b8:4f:ee.
Are you sure you want to continue connecting (yes/no)? yes
Warning: Permanently added '10.136.23.5' (RSA) to the list of known hosts.
User Authentication
Password:
```

sftp>

2．通过SFTP命令进行文件操作

当SFTP客户端登录到SSH服务器之后，用户可以在SFTP客户端进行如表3-31所示的文件操作。有关文件和目录的操作方法参见本书第2章2.4节。但在SFTP客户端视图下，文件操作命令不支持联想功能，必须手动输入完整的命令，否则会提示是不支持的命令 。

表3-31 通过SFTP文件操作命令进行文件操作

命令	说明
cd [*remote-directory*]	改变用户的当前工作目录
cdup	改变用户的工作目录为当前工作目录的上一级目录
pwd	显示用户的当前工作路径
dir/ls [**-l** | **-a**] [*remote-directory*]	显示指定目录下的文件列表。dir 与 ls 执行的效果是一样的
rmdir *remote-directory* &<1-10>	删除服务器上指定的目录。一次最多可以删除 10 个目录。但使用该命令删除目录时，目录中不能有文件，否则会删除失败
mkdir *remote-directory*	在服务器上创建新指定目录
rename *old-name new-name*	改变服务器上指定的文件的名字
get *remote-filename* [*local-filename*]	下载远程服务器上指定的文件
put *local-filename* [*remote-filename*]	上传指定的本地文件到远程服务器
remove *remote-filename* &<1-10>	删除服务器上文件。一次最多可以删除 10 个文件
help [**all** | *command-name*]	请求 SFTP 客户端命令帮助

3．SFTP访问管理

连接成功后，用户可以执行display ssh user-information [username] 命令查看SSH用户信息；执行 display ssh server status命令查看 SSH服务器的全局配置信息；执行display ssh server session命令查看SSH客户端连接会话信息。这几个命令均已在 3.6.4节第5点中有详细介绍，故不再赘述。也可以通过quit命令断开与SFTP服务器的连接。

3.7.5 通过SFTP进行文件操作的配置示例

本示例拓扑结构如图 3-23 所示，终端 PC 与交换机的路由可达， 10.136.23.4 是交换机的管理口 IP 地

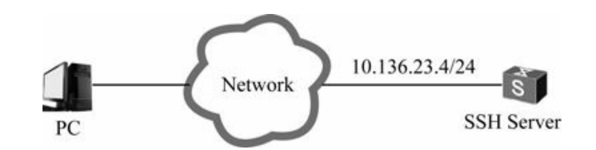


图3-23 通过SFTP进行文件操作的配置示例拓扑结构

址。现希望在SFTP终端与交换机之间 进行安全的文件传输操作，以防止普通FTP服务连接的一些不安全性。现将交换机配置为SSH服务器，提供SFTP服务器功能，通过对客户端的验证和双向的数据加密实现用户对安全文件传输操作的要求。

1．基本配置思路

根据3.7.4节介绍的配置任务及本示例的具体要求，可得出本示例的基本配置思路如下。

（1）在担当SSH服务器的交换机上生成本地密钥对，并使能SFTP服务器功能，实现在服务器端和客户端进行安全的数据交互。

（2）配置用于SFTP连接的VTY用户界面。

（3）配置SSH用户，包括验证方式、服务类型、授权目录以及用户名和密码等。

（4）从终端通过第三方软件OpenSSH实现访问SSH服务器。

2．具体配置步骤

（1）在服务器端生成本地密钥对（在此以RSA加密算法为例），并启用SFTP服务器功能。

<HUAWEI> system-view

[HUAWEI] sysname SSH Server

[SSH Server] rsa local-key-pair create

The key name will be: SSH Server_Host

The range of public key size is (512 ～ 2048).

NOTES: If the key modulus is greater than 512,

        it will take a few minutes.

Input the bits in the modulus[default = 2048] :768

Generating keys. .

......++++++++++++

.........++++++++++++

..++++++++

......++++++++

[SSH Server] sftp server enable

（2）在服务器端配置用于SFTP访问的VTY用户界面（假设为VTY 0～4共五条虚拟通道）。

[SSH Server] user-interface vty 0 4

[SSH Server-ui-vty0-4] authentication-mode aaa

[SSH Server-ui-vty0-4] protocol inbound ssh

[SSH Server-ui-vty0-4] quit

（3）配置SSH用户，包括验证方式、服务类型、授权目录以及用户名和密码等。

[SSH Server] ssh user client001 authentication-type password

[SSH Server] ssh user client001 service-type sftp

[SSH Server] ssh user client001 sftp-directory flash:

[SSH Server] aaa

[SSH Server-aaa] local-user client001 password cipherhuawei@123

[SSH Server-aaa] local-user client001 privilege level 15

[SSH Server-aaa] local-user client001 service-type ssh

[SSH Server-aaa] quit

（4）从终端通过 OpenSSH 软件的 sftp 命令（后面直接接“用户名@交换机的管理IP地址”，以指定登录的用户名和SFTP服务器IP地址）实现访问SFTP服务器，如图3-24 所示。只有在用户终端安装了OpenSSH软件后，Windows命令行提示符才能识别OpenSSH相关命令。

```
C:\WINDOWS\system32\cmd.exe - sftp client001@10.136.23.4

C:\Documents and Settings\Administrator>sftp client001@10.136.23.4
Connecting to 10.136.23.4...
The authenticity of host '10.136.23.4 (10.136.23.4)' can't be established.
RSA key fingerprint is 69:1c:c6:20:5b:29:0e:15:47:50:4f:31:ae:68:5b:0e.
Are you sure you want to continue connecting (yes/no)? yes
Warning: Permanently added '10.136.23.4' (RSA) to the list of known hosts.
User Authentication
Password:
sftp>
```

图3-24 通过sftp命令访问交换机的界面

通过第三方软件连接交换机后，进入客户端的 SFTP 视图，此时可以使用第三方软件支持的SFTP命令执行一系列文件操作。

3.7.6 通过SCP进行文件操作

SCP也是SSH协议的一部分，是基于SSH协议的远程文件复制技术，包括上传和下载。在配置通过SCP进行文件复制之前，也需要确保终端与交换机之间路由可达，且在终端上已安装支持SCP的SSH客户端软件。

1．配置任务

通过SCP进行文件操作的配置任务如下（第1～3步之间没有严格的配置顺序）：

（1）配置SCP服务器功能及参数：包括服务器本地密钥对生成、SCP服务器功能的使能及服务器参数的配置：监听端口号、密钥对更新时间、SSH验证超时时间、SSH验证重试次数等。

因为SCP与STelnet一样都是使用SSH服务，所以本步配置与本章前面3.6.4节的第2点介绍的STelnet服务器功能和参数配置差不多，唯一不同的是在表3-16中的第3步，这里要启用的是SCP服务，使用的命令是 scp server enable，其他配置完全一样，其他配置参见3-16即可。

（2）配置SSH用户登录的用户界面：包括VTY用户界面的用户验证方式、VTY用户界面支持SSH协议及其他基本属性。

本项配置任务的配置方法参见本章 3.5.1～3.5.4 节（在表 3-10 中第 8 步要通过protocol inbound ssh命令配置对应的VTY用户界面支持SSH服务），验证方式的配置参见本章3.5.5节，建议采用AAA验证方式（也可以采用其他验证方式 ）。

（3）配置SSH用户：包括SSH用户的创建、验证方式、服务方式等。

本项配置任务的配置也与3.6.4节介绍的STelnet登录第3点中的SSH用户配置差不多，不同的是要在表 3-17第 4步中使用 ssh user username service-type all命令配置SSH用户支持SCP服务，缺省情况下，SSH用户的服务方式是空，即不支持任何服务方式。其他配置参见表3-16。

在SSH用户验证方面：

① 如果要对SSH用户进行 password、password-dsa或 password-rsa验证，按表 3-18进行配置。

② 如果要对SSH用户进行 dsa、rsa、password-dsa或 password-rsa验证，按表 3-19进行配置。

给与SCP交换机访问的相关参数缺省配置参如表3-32所示。

表3-32 与SCP交换机访问的相关参数缺省配置

参数	缺省值
SCP 服务器功能	关闭
监听端口号	22
服务器密钥对更新时间	0，表示永不更新
SSH 验证超时时间	60s
SSH 验证重试次数	3
SSH 用户	没有创建
SSH 用户的服务方式	空，即不支持任何服务方式

下面仅介绍用户通过SCP协议访问交换机成功后进行文件上传或下载的操作，以及SCP交换机访问管理。

2．通过SCP协议访问交换机进行文件上传或下载操作

从终端通过SCP方式上传或下载文件，需要在终端上安装支持SCP的SSH客户端软件。此处以使用第三方软件OpenSSH和Windows命令行提示符为例进行配置。只有安装了OpenSSH软件后，Windows命令行提示符才能识别OpenSSH相关命令。

可在终端的命令提示符下执行 scp [-portport-number | -a sourceaddress | -i interface- type interface-number | -r | -cipher {des | 3des | aes128 } | -c] * sourcefile destinationfile命令直接上传文件至服务器或从服务器下载文件至本地。当然也可把本地交换机作为SCP客户端与其他配置作为SCP服务器的交换机连接。命令中的参数和选项说明如下。

（1）port-number：可多选参数，指定远端SCP服务器的端口号，取值范围是1～65 535的整数，具体要根据交换机上SCP服务器端口设置而定，缺省为22号TCP端口。

（2）sourceaddress：可多选参数，指定本终端的IP地址。

（3）interface-type interface-number：可多选参数，仅当从本地交换机上连接其他交换机时选用，用于本地交换机SCP连接SCP服务器交换机的源接口。

（4）-r：可多选项，指定进行批量文件上传或下载。

（5）des：多选一选项，指定采用DES算法进行文件传输加密。

（6）3des：多选一选项，指定采用3DES算法进行文件传输加密。

（7）aes128：多选一选项，指定采用AES128算法进行文件传输加密。

（8）-c：可多选项，指定文件上传或下载时进行压缩操作。

（9）sourcefile：指定上传或下载的源文件，文件格式为username@hostname:[path] [filename] 。

（10）destinationfile：指定上传或下载的目标文件，文件格式也为username@hostname:[path] [filename] 。

此处仅以从交换机下载文件到本地终端为例进行介绍。进入Windows的命令行提示符，执行以下OpenSSH命令，按系统提示正确输入用户名和密码，按下回车键后即可开始文件下载。

```
C:\Documents and Settings\Administrator> scp scpuser@10.136.23.5:flash:/vrpcfg.zip vrpcfg-backup.zip
The authenticity of host '10.136.23.5 (10.136.23.5)' can't be established.
RSA key fingerprint is 46:b2:8a:52:88:42:41:d4:af:8f:4a:41:d9:b8:4f:ee.
Are you sure you want to continue connecting (yes/no)? yes
Warning: Permanently added '10.136.23.5' (RSA) to the list of known hosts.
User Authentication
Password:
vrpcfg.zip 100% 1257 1.2KB/s 00:00
Received disconnect from 10.136.23.5: 2: The connection is closed by SSH server
C:\Documents and Settings\Administrator>
```

3．SCP访问管理

用户还可在成功访问交换机后使用display scp-client任意视图命令在SCP客户端查看源配置信息；使用display ssh user-information [username] 任意视图命令在SSH服务器端查看SSH用户信息；使用display ssh server status任意视图命令查看SSH服务器的全局配置信息；使用display ssh server session任意视图命令在SSH服务器端查看SSH客户端连接会话信息。

3.7.7 通过FTPS进行文件操作

FTPS将FTP协议和SSL协议结合，通过SSL对服务器进行验证，对传输的数据进行加密，从而实现了安

全的文件管理操作。在配置通过 FTPS 进行文件操作之前，需要确保终端与交换机之间路由可达，且在终端上已经安装支持SSL的FTP客户端软件。

1．配置任务

通过FTPS进行文件操作的配置任务如下（序号1、2、3、4配置任务中，加载数字证书（序号2）前必须先上传数字证书（序号1），其他无严格配置顺序）。

（1）上传服务器数字证书文件及私钥文件：通过其他文件上传方式将数字证书文件和私钥文件上传至交换机。这项配置任务可以通过前面介绍的FTP、SFTP或者SCP协议进行，不再介绍。

（2）配置SSL策略并加载数字证书：包括配置SSL策略及在服务器上加载数字证书。本项配置任务与3.6.8 节介绍的 HTTPS 登录方式中的 SSL 策略配置和加载方法完全一样，参见表3-24。

（3）配置FTPS服务器功能及FTP服务参数：包括为FTPS服务器配置SSL策略、FTPS服务器的使能及FTP服务参数的配置：端口号、源地址、超时断连时间。本项配置任务与3.7.2节中的表3-27的配置基本一样，将在下面具体介绍。

（4）配置FTP本地用户：包括配置本地用户的服务类型及FTP用户的授权目录。本项配置任务与3.7.2节第3点的配置完全一样，参见其中的表3-28。

（5）用户通过FTPS访问交换机：从终端通过FTPS访问交换机。

通过FTPS访问交换机的相关参数缺省配置如表3-33所示。

表3-33 通过FTPS访问交换机的相关参数缺省配置

参数	缺省值
SSL 策略	没有为 FTPS 服务创建 SSL 策略
FTPS 服务器功能	关闭
监听端口号	21
FTP 用户	没有创建本地用户

下面具体介绍以上配置任务中的第3项和第5项。

2．配置FTPS服务器功能及FTP服务参数

基于FTP协议的FTPS，除了配置FTPS服务器功能外，还可以对FTP服务参数进行配置。具体如表3-34所示，与3.7.2节第2点介绍的FTP交换机访问方式中的FTP服务器功能和参数配置方法类似。

表3-34 FTPS服务器功能及FTP服务参数的配置步骤

步骤	命令	说明
1	**system-view**	进入系统视图
2	**ftp server port** *port-number* 例如：[HUAWEI] **ftp server port** 1028	（可选）指定 FTP 服务器端口号，取值范围为 21 或 1 025～55 535 的整数。缺省情况下，FTP 服务器端监听端口号是 21
3	**ftp secure-server ssl-policy** *policy-name* 例如：[HUAWEI] **ftp secure-server ssl-policy** ftp_server	为 FTPS 服务器配置 SSL 策略，长度范围为 1～23 个字符，不区分大小写，不支持空格。此处配置的 SSL 策略即为前面的配置任务中创建的 SSL 策略。 缺省情况下，FTP 服务器未配置 SSL 策略，可使用 **undo ftp secure-server ssl-policy** 命令删除 FTP 服务器配置的指定 SSL 策略
4	**undo ftp server enable** 例如：[HUAWEI]**undo ftp server enable**	（可选）去使普通 FTP 服务器功能。缺省情况下，交换机的普通 FTP 服务器功能就是关闭的
5	**ftp secure-server enable** 例如：[HUAWEI] **ftp secure-server enable**	使能 FTPS 服务器功能。缺省情况下，未使能 FTPS 服务器。使能 FTPS 服务功能前，必须去使能普通 FTP 服务器功能。
6	**ftp server-source** { **-a** *source-ip-address* | **-i** *interface-type interface- num* } 例如：[HUAWEI]**ftp server- source -i** loopback0	（可选）指定 FTP 服务器的源地址或源接口，实现对交换机进出报文的过滤，保证安全性。二选一参数 *source-ip-address* 用来指定 FTP 服务器源 IP 地址，二选一参数 *interface-type interface-num* 用来指定 FTP 服务器的源接口。但 FTP 服务器端指定的源地址**只能是交换机的 LoopBack 接口 IP 地址或 LoopBack 接口**。配置了服务器的源地址后，登录服务器时所输入的服务器地址必须与该命令中配置的一致，否则无法成功登录。如果在配置此命令前，FTP 服务已经使能，在则在配置本命令后 FTP 服务将重新启动 缺省情况下，FTP 服务器发送报文的源地址为 0.0.0.0，可用 **undo ftp server-source** 命令恢复 FTP 服务器发送报文的源地址为缺省值
7	**ftp timeout** *minutes* 例如：[HUAWEI] **ftp timeout** 20	（可选）配置 FTP 连接最大空闲等待时间，取值范围为 1～35 791 的整数分钟。缺省情况下，连接空闲时间为 30min，可用 **undo ftp timeout** 命令恢复缺省的连接空闲时间 【说明】用户登录到 FTP 服务器后如果连接异常中断或用户非正常中断连接，FTP 服务器是无法知道的，因而连接仍保持着。为防止这类情况的发生，使用连接空闲时间，当连接在一定时间内没有进行命令交互，FTP 服务器即可认为连接已经失效，而断开连接

3．用户通过FTPS访问交换机

需要在用户终端安装支持SSL的FTP客户端软件（如Cuteftp Pro和FlashFXP），通过第三方软件从用户终端登录 FTPS 服务器，实现对 FTPS 服务器文件的安全管理。在此不作具体介绍，参见相关软件的帮助说明。

4．FTPS交换机访问管理

在交换机端可以使用display ssl policy任意视图命令查看配置的SSL策略及加载的数字证书；使用display ftp-server任意视图命令查看FTPS服务器的状态；使用display ftp-users任意视图命令查看登录的FTP用户信息。

3.7.8 通过FTPS进行文件操作的配置示例

本示例拓扑结构如图3-25所示，终端与交换机之间的路由可达，10.137.217.201是 交换机的管理口IP地址。现希望在终端与交换机之间进行安全的文件传输操作，在交换机上部署 SSL 策略，利用数据加密、身份验证和消息完整性验证机制，为网络上数据的传输提供安全性保证。

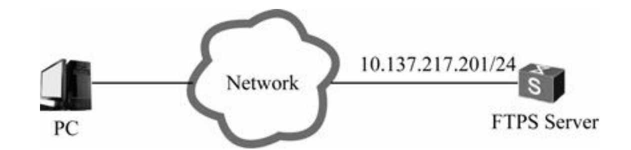


图3-25 通过FTPS进行文件操作的配置示例拓扑结构

1．基本配置思路

根据3.7.7节介绍的配置任务，以及本示例的具体要求可以得出如下的基本配置思路：

（1）配置交换机的普通FTP服务器功能，将PC上存储的数字证书上传到交换机上。然后将位于交换机存储器根目录下的数字证书复制到security目录中。

（2）配置SSL策略并加载数字证书，以实现客户端对服务器的身份验证。

（3）使能FTPS服务器功能，并配置FTP本地用户。

（4）通过终端第三方软件连接FTPS服务器。

2．具体配置步骤

（1）配置服务器的普通FTP服务器功能，配置FTP用户信息（用户名为admin，密码为huawei@123。这里创建的用户可同时作为FTPS用户，因为FTPS所用的也是FTP用户）。

```
<HUAWEI>system-view
[HUAWEI] sysname FTPS-Server
[FTPS-Server] ftp server enable
[FTPS-Server] aaa
[FTPS-Server-aaa] local-user admin password cipherhuawei@123
[FTPS-Server-aaa] local-user admin service-type ftp
[FTPS-Server-aaa] local-user admin privilege level 3
[FTPS-Server-aaa] local-user admin ftp-directory flash:
[FTPS-Server-aaa] quit
[FTPS-Server] quit
```

（2）在终端PC上进入Windows系统命令行提示符输入ftp 10.137.217.201命令，在提示信息中输入正确的用户名和密码与FTP服务器建立FTP连接。然后利用put命令在用户终端将数字证书及私钥文件上传到服务器上，参见3.7.3节介绍的通过FTP协议上传系统文件的配置示例。

上述步骤成功执行后可在交换机执行 dir 命令，此时应该可以看到成功上传的数字证书及私钥文件。如下输出信息中的粗体字部分。

```
<FTPS-Server>dir
Directory of flash:/
  Idx   Attr   Size(Byte)   Date   Time   FileName
  0   drw-   -   May 10 2011 05:05:40   src
  1   -rw-   524,575   May 10 2011 05:05:53   private-data.txt
  2   -rw-   446   May 10 2011 05:05:51   vrpcfg.zip
  3   -rw-   1,302   May 10 2011 05:32:05   servercert.der
  4   -rw-   951   May 10 2011 05:32:44   serverkey.der
..
65,233 KB total (7,289 KB free)
```

（3）配置SSL策略并加载数字证书。在交换机上利用mkdir命令创建security目录， 并利用move命令将位于存储器根目录中的安全证书和密钥文件移动到security目录中。

```
<FTPS-Server>mkdir security/
<FTPS-Server>move servercert.der security/
<FTPS-Server>move serverkey.der security/
```

上述步骤成功执行后可在security目录下执行dir命令，此时应该可看到移动成功的数字证书及私钥文件。如下输出信息中的粗体字部分。

<FTPS-Server>cd security/

<FTPS-Server>dir

Directory of flash:/security/

Idx Attr Size(Byte) Date Time FileName

0 -rw- 1,302 May 10 2011 05:44:34 servercert.der

1 -rw- 951 May 10 2011 05:45:22 serverkey.der

65,233 KB total (7,289 KB free)

（4）创建SSL策略，并加载ASN1格式的数字证书，以确保进行数据传输时的安全性。

<FTPS-Server>system-view

[FTPS-Server] ssl policy ftp_server

[FTPS-Server-ssl-policy-ftp_server] certificate load asn1-cert servercert.der key-pair rsa key-file serverkey.der

[FTPS-Server-ssl-policy-ftp_server] quit

（5）使能FTPS服务器功能，加载SSL策略，并配置FTP本地用户，但FTP用户的配置在前面进行 FTP 文件传输时已创建好，参见上面的第（1）步。另外要注意，使能FTPS服务器功能前，必须先去使能普通FTP服务器功能。

[FTPS-Server] undo ftp server

[FTPS-Server] ftp secure-server ssl-policy ftp_server

[FTPS-Server] ftp secure-server enable

在交换机端执行display ssl policy命令可以看到加载的证书详细信息。具体如下：

[FTPS-Server] display ssl policy

SSL Policy Name: ftp_server

Policy Applicants:

Key-pair Type: RSA

Certificate File Type: ASN1

Certificate Type: certificate

Certificate Filename: servercert.der

Key-file Filename: serverkey.der

Auth-code:

MAC:

CRL File:

Trusted-CA File:

还可在交换机端执行display ftp-server命令查看SSL策略名称、FTPS服务器的状态。具体如下（显示当前FTPS服务器处于运行状态）：

[FTPS-Server] display ftp-server

FTP server is stopped

Max user number 5

User count 1

Timeout value(in minute) 30

Listening port 21

Acl number 0

FTP server's source address   0.0.0.0

FTP SSL policy   ftp_server

FTP Secure-server is running

完成以上配置后，用户可以通过支持SSL的FTP客户端软件与安全FTP服务器建立连接，并实现文件的上传和下载。具体操作过程请参见第三方软件的帮助文档。

第4章　接口及以太网链路配置与管理

本章介绍的是华为S系列交换机的物理以太网端口、各种逻辑接口（如Loopback接口、NULL接口、以太网子接口、Eth-Trunk接口/子接口等）、端口隔离，以及以太网链路聚合（Eth-Trunk）的配置与管理方法。本章内容虽然很基础，但是却是使用率最高的交换机配置之一，特别是以太网链路聚合。

在华为S系列交换机中，端口隔离可以实现在同一VLAN内部的不同端口间隔离。它有两种方式，一种是一个以太网端口与其他以太网端口间的单向隔离，即使其他以太网端口不能与指定以太网端口进行二层通信，实现二层隔离；另一种把需要相互隔离的以太网端口放进一个隔离组中，实现在隔离组中的以太网端口间的二层，甚至三层隔离。这在对同一VLAN中个别以太网端口进行单向或双向隔离时特别需要。

华为交换机中的以太网链路聚合称之为Eth-Trunk，用于交换机设备间的连接。它有“手工负载分担”和“静态LACP”两种工作模式，手工负载分担模式可以实现提升单链路带宽和负载分担功能，静态LACP模式除此之外还可以实现链路备份，这在交换设备互联中应用非常广，也非常实用。

4.1 交换机接口及基础配置

不同的华为 S 系列交换机上可以使用、配置的接口类型不完全一样，如在二层的S2700系列交换机中基本上只有物理的以太网接口、Console接口、管理以太网接口以及Eth-Trunk子接口等二层接口，支持有限的VLANIF接口数量；在S3700及其他三层交换机上，还可以像以太网子接口、VLANIF接口、Loopback接口、Null接口之类的逻辑三层接口。本节首先了解S系列交换机的接口分类和接口编号规则。

4.1.1 接口分类

接口是交换机与网络中的其他设备交换数据的组件，在S系列交换机上一般分为管理接口、物理接口和逻辑接口三类。

1．管理接口

管理接口主要为用户提供配置管理支持，也就是用户通过此类接口可以登录到交换机，并进行配置和管理操作。在华为S系列交换机中除S1700/2700/3700系列外，其他系列均提供Console和MEth（标识为MEth 0/0/1）两种管理接口。管理接口不承担业务传输。

2．物理接口

物理接口是真实存在、有器件支持的接口。物理接口需要承担业务传输。在交换机上一般主要是以太网接口（如百兆以太网接口、千兆以太网接口和万兆以太网接口等）。物理接口又分电口（以双绞线作为传输介质的以太网接口）和光口（以光纤作为传输介质的以太网接口）。

3．逻辑接口

逻辑接口是指能够实现数据交换功能但物理上不存在，需要通过配置建立的接口。如Loopback接口、Null接口、VLANIF接口、Tunnel接口、以太网子接口、Eth-Trunk接口等。逻辑接口需要承担业务传输。这些逻辑接口的作用及配置方法在本章后面有详细介绍。

4.1.2 物理接口编号规则

在华为S系列交换机上的物理接口的编号规则如下（这是指定接口的依据 ）。

（1）非堆叠情况下，交换机采用“槽位号/子卡号/接口序号”的编号规则来定义物理接口。此时，槽位号表示当前交换机的槽位，取值为 0；子卡号表示业务接口板支持的子卡号（对于不支持业务接口板的交换机，子卡号也固定为0）；接口序号表示交换机上各接口的编排顺序号。

（2）堆叠情况下，交换机采用“堆叠号/子卡号/接口序号”的编号规则来定义物理接口。此时，堆叠号表示堆叠 ID，取值为 0～8；子卡号表示业务接口板支持的子卡号； 接口序号表示交换机上各接口的编排顺序号。

S系列交换机上一般有两排业务接口，左下接口从1起始编号，依据先从下到上，再从左到右的规则依次递增编号，如图4-1所示。例如，左上第一个接口编号为0/0/2。

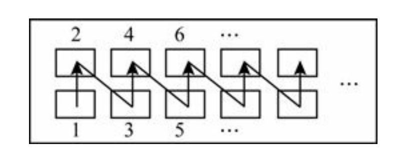


图4-1 物理接口编号规则

4.1.3 接口基本参数配置

本节主要介绍华为S系列交换机接口的基本参数配置，包括接口描述信息、接口流量统计时间间隔功能以及开启或关闭接口。但这些参数配置均是可选配置，因为这些参数都有缺省值 。

为了方便管理和维护设备可以配置接口的描述信息，描述接口所属的设备、接口类型和对端网元设备等信息，以便更好地区分。例如，“设备A连接到设备B 的GE0/0/1接口”可以描述为：To-[DeviceB] GE-0/0/1。

根据配置接口的流量统计时间间隔功能，可以对感兴趣的报文进行统计与分析。流量统计时间间隔的设置既可在系统视图下配置，又可在具体接口视图下配置。在系统视图下配置的接口流量统计时间间隔对时间间隔配置为缺省值的所有接口生效；在接口视图下配置的接口流量统计时间间隔只对本接口生效，不影响其他接口。但在接口视图下配置的时间间隔的优先级高于在系统视图下配置的时间间隔。

接口基本参数的配置方法如表4-1所示（没有严格的先后顺序，其中的序号仅为方便说明）。

表4-1 交换机接口基本参数配置

步骤	命令	说明
1	**system-view** 例如：< HUAWEI > **system-view**	进入系统视图
2	**set flow-stat interval** *interval-time* 例如：[HUAWEI] **set flow-stat interval** 400	（可选）全局配置接口的流量统计时间间隔，取值范围为 10～600 的整数秒，**取值必须是 10 的整数倍** 缺省值是 300s，可用 **undo set flow-stat interval** 命令恢复为缺省值
3	**interface** *interface-type interface-number* 例如: [HUAWEI] **interface Ethernet 0/0/1**	（可选）键入要配置接口基本参数的接口，进入接口视图
4	**description** *description* 例如: [HUAWEI-GigabitEthernet0/0/1] **description** S2700 GigabitEthernet **0/0/1**	（可选）设置接口的描述信息，为 1～242 个字符，支持空格，**区分大小写**。描述信息以输入的第一个非空格字符作为第一个字符开始显示。缺省情况下，接口描述信息为"HUAWEI, HUAWEI Series, *interface-type interface-number* Interface"，可用 **undo description** 命令恢复缺省描述
5	**set flow-stat interval** *interval-time* 例如: [HUAWEI-GigabitEthernet0/0/1] **set flow-stat interval** 400	（可选）为以上接口配置流量统计时间间隔，其他说明参见上面的第 2 步。当在系统视图下的流量统计时间间隔配置与具体接口上的配置不一致时，以此处在接口视图下的配置为准
6	**shutdown** 例如: [HUAWEI-GigabitEthernet0/0/1] **shutdown** **undo shutdown** 例如: [HUAWEI-GigabitEthernet0/0/1] **undo shutdown**	关闭或开启以上接口，缺省情况下，接口处于开启状态。**当修改了接口的工作参数配置，新的配置不能立即生效**，可以依次执行 **shutdown** 和 **undo shutdown** 命令或 **restart** 命令关闭和重启接口，使新的配置生效 但是 Null 接口一直处于 Up 状态，不能使用命令关闭或开启 Null 接口；Loopback 接口一旦被创建，也将一直保持 Up 状态，也不能使用命令关闭或开启

4.1.4 接口配置管理

在日常的交换机管理中，特别是在发生交换机配置故障时，经常要查看接口上的各方面配置信息和运行状态，这时可在用户视图下利用以下一系列 display interface命令或 reset 命令查看或清除相关统计信息。接口统计信息有助于分析接口的故障原因和接口的工作状态，但当需要统计一定时间内接口的流量信息时，需要在统计开始前清除该接口下原有的统计信息。

（1）使用display interface [interface-type [interface-number]] 命令查看所有或具体接口的当前运行状态和统计信息，包括接口当前运行状态、接口基本配置和报文通过接口的转发情况。

（2）使用 display interface brief命令查看各接口状态和配置的简要信息，包括接口的物理状态、协议状态、接收方向最近一段时间的带宽利用率、发送方向最近一段时间的带宽利用率、接收的错误报文数和发送的错误报文数。当要监控接口的状态或检查接口的故障原因时，可执行此命令，根据这些信息进行接口的故障诊断等。

（3）使用display ip interface [interface-type interface-number] 命令查看接口与 IP相关的配置和统计信息，包括接口接收和发送的报文数、字节数和组播报文数，以及接口接收、发送、转发和丢弃的广播报文数。

（4）使用display ip interface brief [interface-type [interface-number]] 命令查看接口与IP相关的摘要信息，包括IP地址、子网掩码、物理链路和协议的Up/Down状态以及处于不同状态的接口数目。在华为 S 系列交换机中，只有像 VLANIF 接口、Loopback接口、以太网子接口等逻辑接口才可以配置IP地址，所以也只能查看这类接口的IP相关的配置和统计信息。

（5）使用display interface description [interface-type [interface-number]] 命令查看指定或所有接口的描述信息。

（6）使用display counters [inbound | outbound] [interface interface-type [interface- number]] 命令查看接口的流量统计。当需要关注接口流量统计的时候可以执行本命令按接口类型或槽位信息查看接口入方向或

出方向的流量统计计数，以便进行故障的定位与排查。

（7）使用 display counters rate [inbound | outbound] [interface interface-type [interface-number]] 命令查看接口的入方向或出方向流量速率。当需要按接口类型关注接口流量速率的时候，用户可以执行本命令按接口类型或槽位信息查看接口入方向或出方向的流量速率，以便于用户进行故障的定位与排查。

（8）使用 reset counters interface [interface-type [interface-number]] 命令清除指定接口的流量统计信息。如果需要统计接口在一段时间内的流量信息，必须在统计开始前使用本命令清除它原有的统计信息，使它重新进行统计。

（9）使用 reset counters if-mib interface [interface-type [interface-number]] 命令清除指定或者所有接口的流量（如Web流量和SNMP流量）统计信息。但执行本命令对 display interface命令显示的接口流量统计信息不会影响。当需要清除 display interface命令查看到的接口统计信息时，可以执行前面介绍的reset counters interface命令。

4.2 以太网接口属性

为了适应网络需求，在华为S系列交换机上定义了以下两种以太网接口类型。

1．二层以太网接口

二层以太网接口是一种物理接口，工作在数据链路层，不能配置IP地址。它可以对接收到的报文进行二层交换转发，可以通过配置各种 Access、Hybrid、Trunk 和 Tunnel这四种端口类型加入一个或多个 VLAN 中，但只能通过三层的 **VLANIF** 接口对接收到的报文进行三层路由转发，所以需要把对应的以太网接口加入到一个VLAN中，然后为该VLANIF接口配置IP地址。

2．三层以太网子接口

三层以太网子接口是一种逻辑接口，工作在网络层，可以配置IP地址，处理三层协议，封装和终结一个或多个VLAN，主要用来实现在三层以太网子接口上收发VLAN报文。用户可以在一个以太网接口上配置多个子接口，这样可对来自不同VLAN的报文从不同的子接口进行转发，为用户提供了很高的灵活性。

注意

当前版本华为S系列交换机不支持三层以太网物理接口，本书中若不做特殊说明， 所指的以太网接口均为二层以太网接口。而且，仅S7700、S9300和S9700三大系列中的 E 系列以太网接口板和 F 系列以太网接口板支持三层以太网子接口，其他系列及S7700、S9300和S9700三大系列中的主控板均不支持。

4.2.1 以太网接口特性

在以太网接口中，可以配置的特性主要包括端口组、接口速率、双模式、自动协商模式支持、流量控制和流量控制自动协商支持、环回测试、电缆检测、端口隔离等方面。下面对这些特性进行简单介绍。

1．端口组

华为S系列交换机支持端口组功能，可方便用户同时对端口组中的多个端口进行一次性配置。这样只需在端口组视图下输入一次配置命令，就可以使该端口组内的所有端口都应用该功能配置，可大大地减轻重复配置工作量。

2．自协商

自协商的主要功能就是使物理链路两端的端口通过协商能力信息交互来自动选择同样的工作参数。自协商端口发送本端的协商能力信息并检测对端的信息，一旦本端收到对端的协商能力信息，并且得知对端也收到本端发送的协商信息时，就比较两端的能力来建立起双方都具有的共同最高性能的工作模式。

自动协商的内容主要包括接入速率和流量控制参数。一旦协商通过，链路两端的端口就以同样的运行速率和流量控制参数来工作。

3．流量控制

流量控制的作用是用来控制发送端的数据发送速率，使接收端设备有能力及时处理来自发送端的数据。当本端和对端设备都开启了流量控制功能后，如果本端交换机发生拥塞，它将向对端设备发送消息，通知对端设备暂时停止发送报文；而对端设备在收到该消息后将暂时停止向本端交换机发送报文，从而避免了报文丢失现象的发生。

4．电缆检测

VCT（Virtual Cable Test，虚拟电缆检测）功能可用于检测接口所连的电缆是否存在故障。当电缆状态正常时显示该电缆的总长度；当电缆状态非正常时显示电缆的故障类型，并且能够给出故障点的位置，便于定位和解决网线问题。

VCT一般用于链路出现故障时检测是否因电缆故障导致的。由于电缆检测会导致业务的短暂中断，所以不建议用户在业务正常运行时使用。

5．环回测试

用户可以开启以太网接口的环回测试功能，检验以太网接口能否正常工作，用以定位芯片内与该接口相关的模块是否出现故障。测试时接口将不能正常转发数据包。S系列交换机支持内部环回测试模式，该测试模式在PHY芯片内部建立自环。端口设置为该模式后会产生一定数量的测试报文，这些报文通过PHY芯片内部建立的自环又返回到该端口。

6．端口隔离

为了实现各接口发送的报文之间二层隔离，可以将不同的以太网端口加入不同的VLAN，但会浪费有限的VLAN资源。采用端口隔离特性，可以实现同一VLAN内端口之间的隔离。用户只需要将端口加入到隔离组中，就可以实现隔离组内端口之间二层数据的隔离。而且这种隔离是双向的，即如果将端口A和B加入同一个隔离组，则从端口A发送的二层报文不能到达端口B，从端口B发送的报文也不能到达端口A。

7．40GE接口拆分

本特性仅S7700、S9300和S9700三大系列交换机支持。它们的40GE接口可以作为一个单独的接口使用，也可以拆分成4个10GE接口。这样40GE接口板可以作为高密度万兆接口板使用，从而增加组网灵活性，减少用户购置成本。

4.2.2 以太网端口组配置与管理

上节介绍了，华为S系列交换机可以配置以太网端口组，实现批量配置，以减少重复配置工作量。这时可以把那些基本属性配置相同的端口加入一个端口组中，然后直接在端口组下配置这些共同的以太网端口属性就可以全面应用到端口组中的所有成员以太网端口上了。

在S2700、S3700T系列交换机中仅支持永久端口组，而在S5700、S6700、S7700、S9300和S9700系列交换机中还支持临时端口组。两种端口组功能相同，不同在于退出临时端口组后，该临时端口组被系统自动删除。

1．配置永久端口组

配置永久端口组的步骤如表4-2所示。

表4-2 永久端口组的配置步骤

步骤	命令	说明
1	**system-view** 例如：< HUAWEI > **system-view**	进入系统视图
2	**port-group** *port-group- name* 例如：[HUAWEI] **port-group** portgroup1	创建并进入永久端口组视图。参数 *port-group-name* 为 1～32 个字符，不支持空格，不区分大小写，但**不能取名为 group** 缺省情况下，系统没有配置永久端口组，可用 **undo port-group** { **all** | *port-group-name* }命令删除指定的或者所有永久端口组
3	**group-member** { *interface-type interface-number1* [**to** *interface-type interface-number2*] } &<1-10> 例如：[HUAWEI-port-group-portgroup1] **group-member** gigabitEthernet0/0/1	将以太网接口添加到指定永久端口组中。命令中的参数说明如下。 （1）*interface-type interface-number1* [**to** *interface-type interface-number2*]：指定添加到永久端口组中的以太网接口。用 **to** 连接的两个接口，表示接口范围是在两个接口编号之间的所有接口（**但必须是相同类型的以太网接口**），如果没有指定 **to** *interface-type interface-number2* 可选参数，则表示添加单个以太网接口 （2）&<1-10>：表示前面的 *interface-type interface-number1* [**to** *interface-type interface-number2*]参数最多可以有 10 个，也就最多可以一次性添加 10 个单个以太网接口，或者 10 个连续编号范围的以太网接口 【注意】在使用 **to** 关键字时要注意以下几个方面。 (1) **to** 关键字前后的**两个接口必须在同一个接口板上**。当有多个接口板的连续接口需要加入时，建议分多次执行该命令或使用多次 **to** 关键字 (2) **to** 关键字前后的**两个接口类型必须相同**，比如同是 Ethernet 接口，或者 GigabitEthernet 接口等 (3) **to** 关键字前后的两个接口必须是具有同一属性的接口，比如同是主接口或同是子接口（仅 S7700、S9300 和 S9700 系列交换机中的 E 系列和 F 系列单板支持以太网子接口）。如果是子接口，**to** 关键字前后的两个子接口**必须是同一个主接口的子接口** 缺省情况下，没有以太网接口添加到永久端口组中，可用 **undo group-member** { *interface-type interface-number1* [**to** *interface-type interface-number2*] } &<1-10>命令删除当前永久端口组中指定端口。在 S5700 和 S6700 系列交换机中还可用 **undo group-member all-unavailable-interface** 命令删除端口组中所有不可用接口
4	**return** 例如：[HUAWEI-port-group-portgroup1] **return**	退出端口组视图，直接返回用户视图
5	**display port-group** [**all** | *port-group-name*] 例如：<HUAWEI> **display port-group**	查看所有或指定永久端口组中的成员接口信息。命令中的参数和选项说明如下： （1）**all**：二选一可选项，指定查看所有端口组及端口组中的成员端口 （2）*port-group-name*：二选一可选参数，指定查看端口成员的端口组名称 如果不配置任何参数，则显示所有永久端口组的名称

【示例 1】配置接口GE0/0/1～GE0/0/3加入端口组portgroup1。

```
<HUAWEI>system-view
[HUAWEI] port-group portgroup1
[HUAWEI-port-group-portgroup1] group-member gigabitethernet 0/0/1 to gigabitethernet 0/0/3
```

【示例 2】查看所有永久端口组及其成员接口信息。从中可以看到交换机上已创建了名为1和2的两个端口组，并且列出了这两个端口组中的以太网端口成员。

```
<HUAWEI>display port-group all
Portgroup: 1
gigabitethernet1/0/1
gigabitethernet1/0/2
gigabitethernet1/0/3
Portgroup: 2
gigabitethernet2/0/1
gigabitethernet2/0/2
gigabitethernet2/0/3
```

2．配置临时端口组

配置临时端口组的方法很简单，只需在系统视图下使用 port-group group-member { interface-type interface-number1 [to interface-type interface-number2] } &<1-10>命令即可。其实它是表4-2中第2步和第3步

命令的结合。具体的参数参见表4-2中的第3步说明。

【示例 3】配置接口GE1/0/1～GE1/0/3加入到临时端口组中。

<HUAWEI>system-view

[HUAWEI] port-group group-membergigabitethernet 1/0/1 to gigabitethernet 1/0/3

4.2.3 以太网接口基本属性配置与管理

以太网接口基本属性包括Combo接口工作模式、接口速率、自协商功能、网线类型、双工模式、流量控制、超大帧支持、能效以太网支持、二/三层模式切换等。它们都可直接在对应的以太网接口视图下进行配置的，但如果有许多交换机端口的以上属性配置一样，也可以先按4.2.2节介绍的端口组，然后在端口组视图下进行批量配置。

说明

Combo接口是一个逻辑接口，一个Combo接口对应设备面板上一个GE电接口和一个GE 光接口，而在设备内部只有一个转发接口。电接口与其对应的光接口是光电复用关系，两者不能同时工作（例如，当激活光接口时，对应的电接口就自动处于禁用状态，反之亦然），用户可根据组网需求选择使用电接口或光接口。电接口和光接口共用一个接口视图。当用户需要激活电接口或光接口、配置电接口或光接口的属性（比如速率、双工模式等）时，在同一接口视图下配置。

以太网接口基本属性的缺省配置如表4-3所示，不同类型以太网在接口速率支持、双工模式、自动协商模式、流量控制和流量控制协商等基本属性的缺省支持如表4-4所示。如果想改变这些缺省配置，可按表4-5所示的方法选择性地进行配置。

表4-3 以太网接口基本属性缺省配置

基本属性	缺省值
Combo 接口工作模式	Auto，即自动切换光口模式与电口模式
MDI 类型	Auto，即自动识别所连接网线的类型
双工模式	自协商模式下，接口的双工模式是与对端协商得到的；非自协商模式下，接口的双工模式为全双工
接口速率	自协商模式下，接口的速率是与对端协商得到的 非自协商模式下，接口的速率为接口支持的最大速率
上报状态变化延时时间	上报 Up 事件延时时间 0ms；上报 Down 事件延时时间 0ms
能效以太网	未使能

表4-4 不同类型以太网接口的基本属性支持

接口类型	速率（Mbit/s）	双工模式	自协商模式	流量控制	流量控制自协商
百兆以太网 FE 电接口	10	全双工/半双工	支持	支持	不支持
	100	全双工/半双工			
千兆以太网 GE 电接口	10	全双工/半双工	支持	支持	支持
	100	全双工/半双工			
	1 000	全双工			
FE 光接口	100	全双工	不支持	支持	不支持
GE 光接口	100	全双工	缺省情况下，GE 光接口不支持自动协商速率，但可使用 **speed auto-negotiation** 命令来配置接口速率自协商功能	支持	支持
	1 000	全双工			
XGE（10GE）光接口	10 000	全双工	不支持	支持	不支持
40GE 光接口	40 000	全双工	不支持	支持	不支持

表4-5 以太网端口的基本属性配置步骤

<table>
<tr><th>配置任务</th><th>命令</th><th colspan="2">说明</th></tr>
<tr><td rowspan="3">公共
配置任务</td><td>system-view
例如：<Sysname> system-view</td><td colspan="2">进入系统视图</td></tr>
<tr><td>interface interface-type interface-number
例如：[HUAWEI] interface gigabitethernet 1/0/1</td><td>进入以太网端口视图</td><td rowspan="2">（二选一），可直接在具体以太网接口视图下个别配置，也可以在对应以太网端口组视图下批量配置，根据实际需要选择</td></tr>
<tr><td>port-group port- group-name
例如：[HUAWEI] port-group portgroup1</td><td>进入端口组视图，在指定的端口组中为各成员端口批量配置基本属性</td></tr>
<tr><td colspan="4">下面是各以太网接口基本属性配置任务</td></tr>
<tr><td>Combo
接口工作
模式配置</td><td>combo-port { auto | copper | fiber }
例如：[HUAWEI-GigabitEthernet1/0/1] combo-port copper
或
[HUAWEI-port-group-portgroup1] combo-port copper</td><td colspan="2">配置 Combo 接口工作模式。命令中的选项说明如下。
• auto：多选一选项，指定自动选择接口模式。如果有光信号则选择光口模式；没有光信号则选择电口模式
• copper：多选一选项，强制选择电口模式，使用双绞线传输数据
• fiber：多选一选项，强制选择光口模式，即使用光纤传输数据
缺省情况下，自动切换光口模式与电口模式，可用 undo combo-port 命令恢复缺省的自动切换模式
【注意】Combo 接口都是 GE 类型接口，所以要在配置 Combo 接口工作模式时，在上一步中必须进入对应的 GE 以太网接口视图</td></tr>
</table>

（续表）

<table>
<tr><th>配置任务</th><th>命令</th><th colspan="2">说明</th></tr>
<tr><td>XGE
以太网
接口工作
模式配置</td><td>set port-work-mode { lan | wan }
例如：[HUAWEI-XGigabitEthernet1/0/1] set port-work-mode wan</td><td colspan="2">配置 XGE 以太网接口的工作模式（仅适用于 S7700、S9300 和 S9700 三大系列中的 XGE 以太网接口）。命令中的选项说明如下。
• lan：二选一选项，指定接口工作在 LAN 模式
• wan：二选一选项，指定接口工作在 WAN 模式
仅可在 XGE 接口视图下配置。缺省情况下，XGE 以太网接口工作在 LAN 模式，可用 undo set port-work-mode 命令恢复 XGE 接口的工作模式为缺省的 LAN 模式</td></tr>
<tr><td>拆分
40GE
接口</td><td>port split
例如：[HUAWEI-40GE1/0/4]port split</td><td colspan="2">将一个 40GE 接口拆分成为 4 个 10GE 接口（仅适用于 S7700、S9300 和 S9700 三大系列中的 40GE 以太网接口），需要在第 2 步进入 40GE 以太网接口视图
缺省情况下，40GE 接口作为一个接口使用，不拆分，可用 undo port split 命令将 4 个 10GE 拆分接口合并成一个 40GE 接口（仅适用于任意一个原来是由 40GE 以太网接口拆分的 10GE 以太网接口），需要先在第 2 步中进入被拆分的 10GE 以太网接口视图</td></tr>
<tr><td rowspan="3">接口
速率配置</td><td>auto speed { 10 | 100 | 1000 }
例如：[HUAWEI-gigabitethernet1/0/1] auto speed 100 1000
或
[HUAWEI-port-group-portgroup1]auto speed 100 1000</td><td colspan="2">在自协商模式下配置以太网接口可协商的接口速率（10Mbit/s、100Mbit/s 或 1 000Mbit/s，可多选），FE 电接口不支持 1 000 这个多选项
缺省情况下，以太网电接口自协商速率范围为接口支持的所有速率，可用 undo auto speed 命令恢复以太网电接口在自协商模式下的协商速率为缺省的接口支持的所有速率</td></tr>
<tr><td>undo negotiation auto
例如：[HUAWEI-gigabitethernet1/0/1] undo negotiation auto
或
[HUAWEI-port-group-portgroup1]undo negotiation auto</td><td rowspan="2">在非自协商模式下配置以太网接口速率</td><td>设置以太网接口工作在非自协商模式下，缺省情况下，以太网接口处于自协商模式，可用 negotiation auto 命令配置以太网接口工作在自协商模式下
【注意】在 GE 光以太网接口下使用 negotiation auto 命令，只能协商双工模式，不能协商速率。如果要在 GE 光接口下启动协商速率功能，则需要使用 speed auto-negotiation 命令。缺省情况下，未使能 GE 光接口速率自协商功能，丁用 undo speed auto-negotiation 命令去使能 GE 光接口速率自协商功能</td></tr>
<tr><td>speed { 10 | 100 | 1000 }
例如：[HUAWEI-gigabitethernet1/0/1] speed 1000
或
[HUAWEI-port-group-portgroup1] speed 1000</td><td>配置以太网接口的接口速率（只能单选，且 FE 电口不支持 1 000 这个选项）
缺省情况下，接口工作于非自协商模式时，它的速率为接口支持的最大速率，可用 undo speed 命令恢复以太网接口在非自协商模式下的速率为缺省值</td></tr>
<tr><td>接口流量
控制配置</td><td>flow-control
例如：[HUAWEI-gigabitethernet1/0/1] flow-control
或
[HUAWEI-port-group-portgroup1] flow-control</td><td colspan="2">配置以太网接口的流量控制功能。对端设备接口也需要打开流量控制开关才能实现流量控制
缺省情况下，未配置以太网接口的流量控制功能，可用 undo flow-control 命令关闭以太网接口的流量控制开关</td></tr>
</table>

（续表）

配置任务	命令	说明	
接口流量控制配置	**negotiation auto** 例如：[HUAWEI-gigabitethernet1/0/1] **negotiation auto** 或 [HUAWEI-port-group-portgroup1] **negotiation auto**	配置以太网接口在自协商模式下的流量控制功能	配置接口工作在自协商模式。缺省情况下，以太网接口处于自协商模式，可用 **undo negotiation auto** 命令配置以太网接口工作在非自协商模式。但 FE 光接口、40GE 光接口不支持配置自协商功能
	flow-control negotiation 例如：[HUAWEI-gigabitethernet1/0/1] **flow-control negotiation** 或 [HUAWEI-port-group-portgroup1]**flow-control negotiation**		配置接口的流量控制自协商功能。对端设备接口也需要配置流量控制自协商功能才能实现流量控制自协商成功，且**只有电接口支持此配置** 缺省情况下，未配置接口的流量控制自协商功能，可用 **undo flow-control negotiation** 命令取消配置以太网接口的流量控制自协商功能
接口双工模式配置	**negotiation auto** 例如：[HUAWEI-gigabitethernet1/0/1] **negotiation auto** 或 [HUAWEI-port-group-portgroup1] **negotiation auto**	配置以太网接口在自协商模式下的双工模式	配置以太网接口工作在自协商模式，同前面介绍
	auto duplex { **full** | **half** }* 例如：[HUAWEI-gigabitethernet1/0/1] **auto duplex half** 或 [HUAWEI-port-group-portgroup1] **auto duplex half**		配置以太网电接口在自协商模式下的双工模式（**full** 代表全双工模式，**half** 代表半双工模式，**可多选**） 仅以太网电接口支持配置双工模式，且 GE 电接口速率为 1 000Mbit/s 时只能为全双工模式，此时如果将双工模式设置为半双工时，接口协商的速率最大为 100Mbit/s。配置以太网电接口为双工模式时，两端接口的双工模式要保持一致 缺省情况下，以太网电接口的双工模式是和对端接口协商得到的，可用 **undo auto duplex** 命令恢复以太网电接口在自协商模式下的双工模式为缺省情况
	undo negotiation auto 例如：[HUAWEI-gigabitethernet1/0/1] **undo negotiation auto** 或 [HUAWEI-port-group-portgroup1]**undo negotiation auto**	配置以太网接口在非自协商模式下的双工模式	配置以太网接口工作在非自协商模式，同前面介绍
	duplex { **full** | **half** } 例如：[HUAWEI-gigabitethernet1/0/1] **duplex full** 或 [HUAWEI-port-group-portgroup1] **duplex full**		配置以太网电接口在非自协商模式下的双工模式（**full** 代表全双工模式，**half** 代表半双工模式）。当希望接口在发送数据包的同时可以接收数据包，可以将接口设置为全双工模式；当希望接口同一时刻只能发送数据包或接收数据包时，可以将接口设置为半双工模式 缺省情况下，当以太网电接口工作在非自动协商模式时，它的双工模式为全双工模式，可用 **undo duplex** 命令用来恢复以太网电接口在非自协商模式下的双工模式为缺省的全双工模式

（续表）

配置任务	命令	说明
执行电缆检测功能	**virtual-cable-test** 例如：[HUAWEI-gigabitethernet1/0/1] **virtual-cable-test**	执行电缆检测功能，对以太网电接口连接电缆进行一次检测，并显示检测的结果。**仅可在具体以太网接口视图下配置**。当电缆状态为正常时，显示信息中的长度是指该电缆的总长度；当电缆状态非正常时，显示信息中的长度是指从本端口到故障位置的长度
接口MDI类型配置	**mdi { across | auto | normal }** 例如：[HUAWEI-gigabitethernet1/0/1] **mdi normal** 或 [HUAWEI-port-group-portgroup1] **mdi normal**	配置以太网电接口 MDI（Medium Dependent Interface，介质相关接口）类型。通过配置以太网电接口 MDI 类型，可以改变引脚在通信中的角色，从而使得接口的网线适应方式与实际使用的网线相匹配。命令中的选项说明如下。 （1）**across**：多选一选项，指定以太网电接口的 MDI 类型为 Across（交叉电缆类型） （2）**normal**：多选一选项，指定以太网电接口的 MDI 类型为 Normal（直通电缆类型） （3）**auto**：多选一选项，指定以太网电接口的 MDI 类型为 Auto（自动识别类型）。自动模式就是自动识别线序，并协商收发的顺序。保证了不管使用何种线序的网线，也不论对端设备是否是同种类型的，都可以正常通信。它的好处就是可以不用考虑双绞线的类型，也不用关心对端设备是否支持 MDI，都能够正常工作 缺省情况下，以太网电接口 MDI 类型也为 Auto 类型，接口能自动识别所连接网线的类型，可用 **undo mdi** 命令恢复以太网电接口 MDI 类型为缺省的自动识别类型 【注意】两台工作于 Across 模式的设备对接，必须使用交叉网线；一端是 Normal 模式，另一端是 Across 模式，则必须使用直通网线。Auto 类型能满足绝大多数的场合，仅当设备不能获取网线类型参数时，需要将模式手工设置为 Across 或 Normal 使用直通网线时，设备两端应该配置不同的类型（如一端为 Across，另一端为 Normal）；使用交叉网线时，设备两端应该配置相同的类型(如同时为 Across 或 Normal，或者至少有一端是 Auto)
接口环回功能使能	**loopback-detect enable** 例如：[HUAWEI-gigabitethernet1/0/1] **loopback-detect enable** 或 [HUAWEI-port-group-portgroup1] **loopback-detect enable**	用来使能接口的环回检测功能，使能本功能后，CPU 会构造检测报文并向链路上发送。如果端口上的发送和接收包一致，则表明此端口的发送和接收链路正常，端口可以正常地进行业务转发 缺省情况下，接口的环回检测功能处于关闭状态，可用 **undo loopback-detect enable** 命令关闭接口的环回检测功能
能效以太网功能使能	**energy-efficient-ethernet enable** 例如：[HUAWEI-gigabitethernet1/0/1] **energy-efficient-ethernet enable** 或 [HUAWEI-port-group-portgroup1]**energy-efficient-ethernet enable**	使能以太网电接口的能效以太网功能（除 S2700、S3700 系列外其他支持 VRP 系统的 Sx700 系列交换机均支持）。使用本命令使能电接口的能效以太网功能后，当接口处于业务空闲状态时系统会自动调节给该接口的供电，当正常传输数据时，则恢复接口正常供电 【注意】能效以太网功能只能在电接口上配置，光接口或 Combo 口和 10Mbit/s 速率的电接口都不支持 缺省情况下，未使能电接口的能效以太网功能，可用 **undo energy-efficient-ethernet enable** 命令去使能电接口的能效以太网功能

（续表）

配置任务	命令	说明
接口最长帧配置	**jumboframe enable** [*value*] 例如：[Sysname-Ethernet1/0/1] **jumboframe enable** 5000 或 [HUAWEI-port-group-portgroup1] **jumboframe enable** 5000	指定允许通过以太网接口的最大帧长。参数 *value* 的取值范围在不同 S 系列交换机可能不一样，如 S5700EI 系列为 1 600～9 712 的整数，S5700SI/5710LI 系列为 1 600～10 224 的整数，S5700LI 系列为 1 536～10 240 的整数，S6700 系列为 1 536～12 288 的整数，S3700EI/3700SI 系列为 1 600～13 296 的整数，其他系列为 1 536～12 288 的整数，单位是字节 不同 S 系列交换机允许通过的最大帧长的缺省值也可能不一样，如 S3700EI/3700SI 系列为 1 600 字节，S5700EI/5700SI/5710LI 系列为 1 600 字节，其他系列为 9 216 字节，可用 **undo jumboframe enable** 命令恢复接口允许通过的最大帧长为缺省值
二/三层模式切换	**undo portswitch** 例如：[Sysname-Ethernet1/0/1]**undo portswitch** 或 [HUAWEI-port-group-portgroup1] **undo portswitch**	将以太网接口从二层模式切换到三层模式（**除 S2700 和 S3700 系列外，其他支持 VRP 系统的 Sx700 系列均支持**）。还可通过 **undo portswitch batch** *interface-type* { *interface-number1* [**to** *interface-number2*] } <1-10>命令批量配置接口切换到三层模式（**S2700、S3700 和 S9300 系列不支持该命令**） 当需要使用某个交换机以太网接口作在 PE（提供商边缘）与 CE（客户端边缘）相连的接口时，需要使用本命令将二层口切换为三层口。但一定要注意，在华为 S 系列交换机中，**工作在三层模式的以太网接口也不能配置 IP 地址** 缺省情况，以太网接口工作在二层模式，可用 **portswitch** 或 **portswitch batch** *interface-type* { *interface-number1* [**to** *interface-number2*] } <1-10>（**S2700、S3700 和 S9300 系列不支持该命令**）命令将单个或批量以太网接口从三层模式切换到二层模式

【示例 1】将一个40GE接口拆分为四个10GE接口，并重启接口板使配置生效。

<HUAWEI> system view

[HUAWEI] interface40GE 1/0/4

[HUAWEI-40GE1/0/4] port split

Warning: This command will take effect only after resetting the board. 40GE2/2/0

/0 will not be changed. 40GE2/2/0/1-40GE2/2/0/7 will be split up into XGE, and t

he port configuration will be lost when the port type is changed! Continue? [Y/N

] :y

Info: Succeeded in setting the configuration.

[HUAWEI-40GE1/0/4] quit

[HUAWEI] quit

<HUAWEI> reset slot 1

Caution!!! Confirm to reset slot 1 ? [Y/N]:y

INFO: The board[1] reset success!

【示例 2】对以太网电口GE1/0/1连接电缆进行检测。在检测结果中显示了4对网线全部正常（状态为OK）。输出信息中的字段说明如表4-6所示。

<HUAWEI> system-view

[HUAWEI] interface gigabitethernet 1/0/1

[HUAWEI-GigabitEthernet1/0/1] virtual-cable-test

Warning: The command will stop service for a while, Continue?[Y/N]y

Pair A length: 189meter(s)

Pair B length: 189meter(s)

Pair C length: 189meter(s)

Pair D length: 189meter(s)

Pair A state: OK

Pair B state: OK

Pair C state: OK

Pair D state: OK

表4-6 virtual-cable-test命令输出信息字段说明

字段	说明
Pair A/B/C/D	表示网线的 4 对线
Pair A length	表示网线长度：有故障时为端口到故障位置的长度；无故障时为网线的实际长度。未接网线时为缺省长度 0 米
Pair A state	网线状态，包括：Ok（正常）、Open（开路）、Short（短路）、Crosstalk（线序错误）、Unknown（其他未知故障原因）

4.2.4 接口频繁Up/Down故障分析与排除

接口频繁 Up/Down 是我们经常遇到的一个交换机故障，这通常是由于链路两端接口的双工模式、速率、协商模式配置不一致造成的。此时可在交换机上执行 display interface [interface-type [interface-number]] 命令查看对应接口的双工模式、速率、协商模式配置信息，如下所示（注意粗体字部分显示），然后进行下面的分析。

<HUAWEI>display interface ethernet 0/0/1

Ethernet0/0/1 current state : UP

Line protocol current state : UP

Description:HUAWEI, HUAWEI Series, Ethernet0/0/1 Interface

Switch Port,PVID : 1,TPID : 8100(Hex),The Maximum Frame Length is 2044

IP Sending Frames' Format is PKTFMT_ETHNT_2, Hardware address is 0025-9e80-2494

Port Mode: COMMON COPPER

Speed : 100, Loopback: NONE

Duplex: FULL, Negotiation: ENABLE

Mdi : AUTO

Last 300 seconds input rate 760 bits/sec, 0 packets/sec

Last 300 seconds output rate 896 bits/sec, 0 packets/sec

Input peak rate 12304 bits/sec,Record time: 2010-08-05 10:32:18

Output peak rate 14568 bits/sec,Record time: 2010-08-03 08:47:01

Input: 28643 packets, 2734204 bytes

Unicast   :   20923,Multicast   :   7703

Broadcast   :   17,Jumbo   :   0

CRC   :   0,Giants   :   0

Jabbers   :   0,Fragments   :   0

Runts   :   0,DropEvents   :   0

Alignments   :   0,Symbols   :   0

Ignoreds   :   0,Frames   :   0

Discard   :   474,Total Error   :   0

Output: 68604 packets, 8057155 bytes

Unicast   :   20429,Multicast   :   14054

Broadcast   :   34121,Jumbo   :   0

Collisions : 0,Deferreds : 0
Late Collisions : 0,ExcessiveCollisions : 0
Buffers Purged :0
Discard : 0,Total Error : 0
Input bandwidth utilization threshold : 100.00%
Output bandwidth utilization threshold: 100.00%
Input bandwidth utilization : 0.01%
Output bandwidth utilization : 0.00%

（1）先查看输出信息中的Negotiation字段，如果显示为“ENABLE”，则表示该接口工作在自协商状态下；如果显示为“DISABLE”，则表示该接口工作在非自协商状态下。保持两边的协商模式一致，可在接口视图下可以使用 negotiation auto命令启用接口的自协商模式。如果自协商模式下接口仍然频繁Up/Down，可以尝试使用undonegotiation auto命令将接口改成非自协商模式，然后调整两边接口的速率、双工模式一致。

（2）查看输出信息中的Speed字段，在非自协商模式下如果设备两端接口速率不一致，则在接口视图下使用 speed { 10 | 100 | 1000 }命令调整链路两端接口速率一致。

（3）再查看输出信息中的Duplex 字段，在非自协商模式下如果设备两端接口双工模式不一致，则在接口视图下使用duplex { full |half }命令调整链路两端接口双工模式一致。

4.3 端口隔离

以前为了实现报文之间的二、三层隔离，是采用将不同端口加入不同VLAN的方法实现，这样不但配置比较麻烦，而且浪费了有限的VLAN资源。另外，在同一个VLAN中各端口至少是二层互通的，也达不到完全的端口隔离的目的。采用端口隔离特性就可以实现同一VLAN内端口之间的二层隔离，只需要将端口加入到隔离组中，可为用户提供更安全、更灵活的组网方案。但这种方法都仅适用于在同一交换机上不同端口间的隔离，且一个端口可以加入多个端口隔离组 。

4.3.1 端口隔离配置与管理

在华为S系列交换机中，支持“接口单向隔离”和“端口隔离组”这两种端口隔离方法。“接口单向隔离”是在要阻止某个本地端口发送的报文到达其他端口，而不限制其他端口的报文到达本地端口 时所采用的隔离方法。如接口A与接口B之间单向隔离，即接口A发送的报文不能到达接口B，但从接口B发送的报文可以到达接口A。“端口隔离组”是在要实现一组端口间相互（双向）二层隔离时所采用的隔离方法。同一端口隔离组的接口之间互相隔离，不同端口隔离组的接口之间不隔离。下面分别介绍具体的配置方法。

1．配置端口单向隔离

配置端口单向隔离的方法是在具体的以太网接口视图下指定一个或者多个需要对当前接口隔离的其他以太网接口，也就使得当前端口不能发送数据给这些接口，但不限制这些端口发送数据给当前端口 。具体的配置方法如表4-7所示。

表4-7 端口单向隔离的配置步骤

步骤	命令	说明
1	**system-view** 例如：< HUAWEI > **system-view**	进入系统视图
2	**port-isolate mode** { **l2** | **all** } 例如：[HUAWEI] **set flow-stat interval** 400	（可选）全局配置端口隔离模式。命令中的选项说明如下： （1）**l2**：二选一选项，指定端口隔离模式为二层隔离，三层互通 （2）**all**：二选一选项，指定端口隔离模式为二层和三层都隔离。但 S2700 系列中**除 S2700-52P-EI、S2700-52P-PWR-EI 和 S2710-SI 机型外，其他机型均仅支持二层隔离，三层互通** 缺省情况下，端口隔离模式为二层隔离、三层互通，可用 **undo port-isolate mode** 命令恢复端口隔离模式为缺省模式
3	**interface** *interface-type interface-number* 例如：[HUAWEI] **interface Ethernet 0/0/1**	键入要配置端口单向隔离的接口，进入接口视图

（续表）

步骤	命令	说明
4	**am isolate** { *interface-type interface-number* }&<1-8> 或 **am isolate** *interface-type interface-number* [**to** *interface-number*] 例如：[HUAWEI-GigabitEthernet0/0/1]**am isolate** gigabitethernet0/0/2 **to** 0/0/4	配置当前端口与指定端口的单向隔离，参数 *interface-type interface-number* 和 *interface-type interface-number* [**to** *interface-number*]用来指定要与当前端口单向隔离的接口列表，参数&<1-8>用来指定最多可以有 8 个接口或接口列表 【说明】端口单向隔离支持不同类型的端口混合隔离，但不支持端口与管理网口单向隔离，也不支持Eth-Trunk与自身成员端口单向隔离 缺省情况下，未配置端口单向隔离，可用 **undo am isolate** [{ *interface-type interface-number* }&<1-8>]或者 **undo am isolate** [*interface-type interface-number* [**to** *interface-number*]] 命令取消当前端口与指定端口的单向隔离；如果不指定参数表示取消当前端口与所有端口的单向隔离配置

2．配置端口隔离组

配置端口隔离组的方法只需要把想相互隔离的以太网接口加入到同一个隔离组中即可，具体的配置步骤如表4-8所示。

表4-8 端口隔离组的配置步骤

步骤	命令	说明
1	**system-view** 例如：< HUAWEI > **system-view**	进入系统视图
2	**port-isolate mode** { **l2** | **all** } 例如：[HUAWEI] **set flow-stat interval** 400	（可选）全局配置端口隔离模式，其他说明参见表 4-7 中的第 2 步
3	**interface** *interface-type interface-number* 例如：[HUAWEI] **interface** GigabitEthernet 0/0/1	键入要加入端口隔离组的接口，进入接口视图
4	**port-isolate enable** [**group** *group-id*] 例如：[HUAWEI-GigabitEthernet0/0/1] **port-isolate enable group** 10	使能端口隔离功能，并把以上端口加入由可选参数 *group-id* 指定的端口隔离组中（在指定端口隔离组的同时会创建相应的组）。如果不指定 *group-id* 可选参数，则缺省加入的端口隔离组为 1 【注意】要相互隔离的端口一定要加入到同一个端口隔离组，否则不会起到隔离的作用，因为同一端口隔离组的端口之间互相隔离，不同端口隔离组的端口之间不隔离 缺省情况下，未使能端口隔离功能，可用 **undo port-isolate enable** 命令关闭端口的隔离功能
5	**return** 例如：[HUAWEI-GigabitEthernet0/0/1] **return**	退出接口视图，直接返回用户视图
6	**display port-isolate group** { *group-id* | **all** } 例如: <HUAWEI>**display port-isolate group** 10	（可选）查看端口隔离组的成员配置。命令中的参数和选项说明如下： （1）*group-id*：二选一参数，指定要查看隔离组成员配置信息的端口隔离组编号，为 1～64 的整数 （2）**all**：二选一选项，指定查看所有已有的端口隔离组的成员配置信息

【示例 】查看交换机上当前所有端口隔离组的配置。从中可以看出，当前有两个端口隔离组（组号分

别为3和4），并且可以看到它们各自所包括的端口成员。

<HUAWEI>display port-isolate group all

The ports in isolate group 3:

GigabitEthernet1/0/1 GigabitEthernet1/0/2

The ports in isolate group 4:

GigabitEthernet2/0/1 GigabitEthernet2/0/2

4.3.2 端口隔离配置示例

本示例拓扑结构如图4-2所示，PC1、PC2和PC3连接在同一交换机上，同属于VLAN 10，且位于同一IP网段。现希望PC1与PC2之间不能二层互访，PC1与PC3之间以及PC2与PC3之间都可以二层互访。

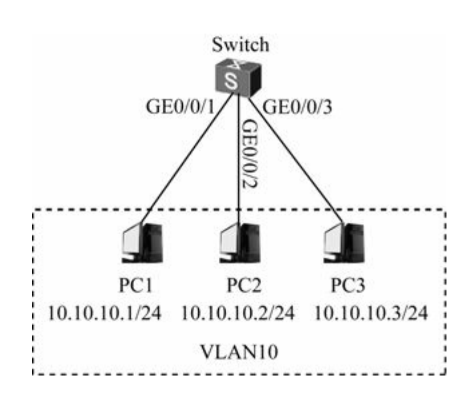


图4-2 端口隔离配置示例拓扑结构

本示例的配置很简单，因为需要PC1和PC2间不能进行二层互访，所以需要采用端口隔离组方法来进行端口隔离。只需要将PC1和PC2所连接的交换机端口（分别为GE0/0/1和GE0/0/2）加入到隔离组中就可以达到目的。

因为S系列交换机的端口隔离模式缺省是二层隔离，三层互通，满足本示例要求，所以不需要另外配置端口隔离模式。下面是具体的配置步骤（VLAN部分的配置不包括在其中）。

（1）配置GE0/0/1的端口隔离组隔离功能（假设加入的端口隔离组号为10）。

<HUAWEI>system-view

[HUAWEI] interfacegigabitethernet 0/0/1

[HUAWEI-GigabitEthernet0/0/1] port-isolate enable group 10

[HUAWEI-GigabitEthernet0/0/1] quit

（2）配置GE0/0/2的端口隔离组隔离功能。注意，此时所加的端口隔离组编号一定要与前面GE0/0/1端口加入的端口隔离组的编号一样。

[HUAWEI] interfacegigabitethernet 0/0/2

[HUAWEI-GigabitEthernet0/0/2] port-isolate enable group 10

[HUAWEI-GigabitEthernet0/0/2] quit

4.4 逻辑接口配置与管理

在华为交换机中，除了那些物理以太网接口外，还有许多用于业务处理的逻辑接口。有关S系列交换机中的主要逻辑接口及其特性说明如表4-9所示。

表4-9 S系列交换机中的主要逻辑接口及特性说明

逻辑接口类型	说明
Eth-Trunk 子接口	是一种具有二层特性和三层特性的逻辑接口，能把多个以太网接口在逻辑上等同于一个逻辑接口（相当于端口聚合），比单个物理以太网接口具有更大的带宽和更高的可靠性。**将在本章后面介绍 Eth-Trunk 链路聚合时介绍**
Tunnel 接口	是一种具有三层特性的逻辑接口，隧道两端的设备利用 Tunnel 接口发送报文、识别并处理来自隧道的报文
VLANIF 接口	是一个具有三层特性的逻辑接口，通过配置 VLANIF 接口的 IP 地址可实现 Vlan 间的互访。将在本书第 6 章介绍

（续表）

逻辑接口类型	说明
以太网子接口	以太网子接口就是在一个主以太网接口上配置的虚拟接口，是一种具有三层特性的逻辑接口，主要用于实现与多个远端进行通信。但目前只有 S5700HI、S5710EI 系列（**但不能配置 IP 地址**），以及 S7700、S9300 和 S9700 系列 E 系列和 F 系列单板（**可以配置 IP 地址**）支持子接口配置
Loopback 接口	是一种具有三层特性的逻辑接口，主要应用其接口状态永远是 Up，并且可以配置 32 位子网掩码的特性
NULL 接口	主要用于路由过滤等特性，因为任何送到该接口的网络数据报文都会被丢弃

下面仅介绍以太网子接口、Loopback接口和NULL接口这三种逻辑接口。

4.4.1 以太网子接口配置与管理

以太网子接口可用于VLAN间的三层互通和局域网与广域网间的互联。在三层互通方面，我们知道VLAN可将一个物理的LAN在逻辑上划分多个广播域，VLAN内的主机可以直接互相二层通信，而 VLAN间的主机不能互相二层通信。要实现不同 VLAN间用户互通必须借用三层技术。

在华为设备中，目前有以下两种方法可实现

（1）在三层交换机上通过VLANIF接口实现。

（2）在路由器（包括S系列中的路由交换机，如S7700、S9300和S9700系列）上通过三层以太网接口实现，即通常所说的单臂路由。

但是传统的三层以太网接口不支持VLAN报文，当收到VLAN报文时会当成非法报文而丢弃。为了实现VLAN间的互通，在三层以太网接口上可创建以太网子接口，通过在子接口上部署终结子接口功能将VLAN报文中的VLAN标签剥离掉，从而实现VLAN间的三层互通。

在局域网和广域网的互联方面，局域网内的报文大多数都带有VLAN标签，但是一些广域网协议并不能识别VLAN报文，如ATM、FR和PPP等。这种情况下如果需要将局域网中的VLAN报文转发到广域网中，则需要在出接口上创建子接口，VLAN报文时先在本地记录报文的VLAN信息，然后剥掉VLAN标签后再转发。

属于不同 VLAN 且位于不同网段的用户，可通过部署子接口、配置 IP 地址（仅**S7700/9300/9300E/9700**系列支持 ）并与VLAN相关联，通过三层网络实现VLAN间通信。以太网子接口的配置步骤如表4-10所示。

表4-10 以太网子接口的配置步骤

步骤	命令	说明
1	**system-view** 例如：<HUAWEI>**system-view**	进入系统视图
2	**interface** *interface-type interface-number.subinterface-number* 例如：[HUAWEI] **interface** GigabitEthernet 0/0/1.1	进入指定以太网子接口的视图（**仅 S5700HI/5710EI 系列、S7700/9300/9300E/9700 系列 E 系列和 F 系列单板支持**）。命令中的参数说明如下。 （1）*interface-type interface-number*：指定要划分子接口的物理以太网接口 （2）*.subinterface-number*：指定创建的子接口编号，取值范围为 1～4 096 整数 **但 Eth-Trunk 中的成员以太网接口上不能创建子接口**

（续表）

<table>
<tr><th>步骤</th><th>命令</th><th colspan="2">说明</th></tr>
<tr><td>3</td><td>ip address ip-address { mask | mask-length } [sub]
例如：[HUAWEI-GigabitEthernet0/0/1.1] ip address 192.168.0.10 255.255.255.0</td><td colspan="2">（可选）为以太网子接口配置 IP 地址（仅 S7700/9300/9300E/9700 系列 E 系列和 F 系列单板支持）。命令中的参数和选项说明如下。
（1）ip-address：指定子接口的 IP 地址
（2）mask：二选一参数，指定子接口 IP 地址对应的子网掩码
（3）mask-length：二选一参数，指定子接口 IP 地址对应的子网掩码长度
（4）sub：可选项，指定所配置的 IP 地址为子接口的从 IP 地址，如果不选择此可选项，则配置的 IP 地址为子接口的主 IP 地址
当为一个以太网子接口配置两个乃至两个以上的 IP 地址时，对第二个及以后的 IP 地址必须用关键字 sub 指示
缺省情况下，在子接口上没有配置 IP 地址，可用 undo ip address ip-address { mask | mask-length } [sub] 删除子接口上指定的 IP 地址</td></tr>
<tr><td rowspan="2">4</td><td>dot1q termination vid low-pe-vid [to high-pe-vid]
例如：[HUAWEI-GigabitEthernet0/0/1.1] dot1q termination vid 4</td><td rowspan="2">（二选一）根据 VLAN 报文携带的 Tag 层数可以将 VLAN 报文分为 Dot1q 报文（带有一层 VLAN Tag）和 QinQ 报文（带有两层 VLAN Tag）。相应的终结也分为两种：Dot1q 终结用来终结 Dot1q 报文，QinQ 终结用来终结 QinQ 报文。您可以根据实际情况选择这两条命令之一来进行配置。接收报文时，剥掉报文中携带的 Tag 后进行三层转发。转发出去的报文是否带 VLAN 标签由出接口决定；发送报文时，将相应的 VLAN 信息添加到报文中再发送
有关 VLAN 及 QinQ 方面的知识分别参见本书第 6、第 7 章</td><td>配置子接口对一层 Tag 报文的终结功能。命令中的参数说明如下。
（1）low-pe-vid：指定用户报文中的 VLAN 标签的取值下限，取值范围为 2～4 094 的整数
（2）high-pe-vid：可选参数，指定用户报文中的 VLAN 标签的取值上限，取值范围为 2～4 094 的整数
子接口收到的用户报文的 VLAN 标签值应该在命令中指定的 pe-vid 范围内，否则该报文将被丢弃。但当子接口用于三层转发时（实现 VLAN 间互通，也就是我们平常所说的单臂路由），不支持将通过的 VLAN 配置成一段
缺省情况下，子接口没有配置 dot1q 封装的单层 VLAN ID，可用 undo dot1q termination vid low-pe-vid [to high-pe-vid] 命令取消子接口 dot1q 封装的单层 VLAN ID</td></tr>
<tr><td>qinq termination pe-vid pe-vid ce-vid ce-vid1 [to ce-vid2]
例如：[HUAWEI-GigabitEthernet0/0/1.1] qinq termination pe-vid 4 ce-vid 10 to 12</td><td>配置子接口对两层 Tag 报文的终结功能。命令中的参数说明如下：
（1）pe-vid：指定 PE 的 VLAN ID，即允许通过的外层 VLAN 标签的值，取值范围为 2～4 094 的整数
（2）ce-vid1：指定 CE 的 VLAN ID，即允许通过的内层 VLAN 标签的值，用户报文内层 VLAN 标签的取值下限，取值范围为 1～4 094 的整数
（3）ce-vid2：可选参数，指定 CE 的 VLAN ID，即允许通过的内层 VLAN 标签的值，用户报文内层 VLAN 标签的取值上限，取值范围为 2～4 094 的整数
子接口收到的用户报文的 VLAN 标签值应该在命令中指定的 PE 和 CE 的 VLAN 标签范围内，否则该报文将被丢弃。但当子接口用于三层转发时（实现 VLAN 间互通，也就是我们平常所说的单臂路由），不支持将通过的 VLAN 配置成一段
缺省情况，子接口没有配置 QinQ 封装的双层 VLAN ID，可用 undo qinq termination pe-vid ce-vid 命令取消子接口 QinQ 封装的双层 VLAN ID</td></tr>
</table>

（续表）

步骤	命令	说明
5	**arp broadcast enable** 例如：[HUAWEI-GigabitEthernet0/0/1.1] **arp broadcast enable**	使能子接口的 ARP 广播功能。因为缺省情况下，终结子接口不能转发 ARP 广播报文，在收到 ARP 广播报文后直接丢弃。这样一来，就无法实现子接口上的三层转发功能。为了允许终结子接口能转发 ARP 广播报文，可以通过在子接口上使用本命令使能终结子接口的 ARP 广播功能。当 IP 报文需要从终结子接口发出，但是在对应主接口上没有目的主机相应的 ARP 表项时： （1）当接入设备能够主动发送 ARP 报文时，则不需要配置终结子接口的 ARP 广播功能，就可以实现从该终结子接口的转发。 （2）当接入设备不能够主动发送 ARP 报文时，如果终结子接口上未使能 ARP 广播功能，那么系统会直接把该 IP 报文丢弃。此时该终结子接口的路由可以看作是黑洞路由。如果终结子接口上已使能 ARP 广播功能，则系统会构造带 VLAN 标签的 ARP 广播报文，然后再从该终结子接口发送出去 使能或去使能子接口的 ARP 广播功能，会使该子接口的路由状态发生一次先 Down 再 Up 的变化，从而可能导致整个网络的路由发生一次震荡，影响正在运行的业务 缺省情况下，终结子接口没有使能 ARP 广播功能 **undo arp broadcast enable** 命令去使能终结子接口的 ARP 广播功能

可用以下命令查看以太网子接口上的相关配置：

（1）display interface [interface-type [interface-number [.subnumber]]]：查看指定或者所有以太网子接口的状态。

（2）display dot1q information termination [interface interface-type interface-number [.subinterface-number]]：查看配置了 dotlq终结的所有接口的名称以及终结子接口对用户报文终结的规则数量。

（3）display qinq information termination [interface interface-type interface-number [.subinterface-number]]：查看配置了QinQ终结的指定或者所有接口的名称以及终结子接口对用户报文终结的规则数量。

【示例 1】在GE1/0/1.1子接口上配置封装 dot1q VLAN 100。

<HUAWEI>system-view

[HUAWEI] interface GigabitEthernet1/0/1.1

[HUAWEI-GigabitEthernet1/0/1.1] dot1q termination vid 100

【示例 2】在GE1/0/1.1子接口上配置封装QinQ：报文外层VLAN ID为 100，内层VLAN ID为 200。

<HUAWEI>system-view

[HUAWEI] interface GigabitEthernet1/0/1.1

[HUAWEI-GigabitEthernet1/0/1.1] qinq termination pe-vid 100 ce-vid 200

【示例 3】查看所有配置dot1q封装方式的子接口信息。输出信息字段说明如表4-11所示。

< HUAWEI >display dot1q information termination

GigabitEthernet0/0/1.3

Total QinQ Num: 1

dot1q termination vid 3

Total vlan-group Num: 0

表4-11 display dot1q information命令输出信息字段说明

字段	说明
GigabitEthernet0/0/1.3	显示子接口的名称
Total QinQ Num	显示子接口对用户报文配置规则的数量
dot1q termination vid 3	显示子接口配置允许通过的 VLAN ID
Total vlan-group Num	显示子接口下配置的用户 VLAN 组的数量

【示例 4】查看配置了 QinQ 封装方式的所有子接口。输出信息字段说明参见表4-11。

< HUAWEI >display qinq information termination

GigabitEthernet0/0/1.30

Total QinQ Num: 1

qinq termination pe-vid 300 ce-vid 200

Total vlan-group Num: 0

4.4.2 Loopback接口配置与管理

Loopback是一种三层逻辑接口，且在一台交换机上可以创建多个Loopback接口。创建Loopback接口后，该接口会一直保持Up状态（但是可以删除），所以用户可通过配置Loopback接口达到提高网络可靠性的目的。而且Loopback接口可以配置32位掩码的IP地址。基于上述特点，Loopback接口通常有以下几种主要应用：

（1）将Loopback接口的IP地址指定为报文的源地址，可以提高网络可靠性。

（2）在一些动态路由协议中，当没有配置Router ID时，将选取所有Loopback接口上数值最大的 IP地址作为Router ID。

（3）在BGP协议中，将发送BGP报文的源接口配置成Loopback接口可以保证BGP会话不受物理接口故障的影响。

（4）Loopback接口可以配置掩码为全1的IP地址，从而可以节约IP地址。

（5）Loopback接口可以配置IPv4地址，可以用于绑定VPN实例、对源IPv4地址进行校验。

Loopback接口只能配置IP地址及报文的源IP地址检查，具体配置步骤如表4-12所示。

表4-12 Loopback接口的配置步骤

步骤	命令	说明
1	**system-view** 例如：< HUAWEI > **system-view**	进入系统视图
2	**interface loopback** *loopback-number* 例如：[HUAWEI] **interface loopback** 1	创建并进入 Loopback 接口视图。参数 *loopback-number* 用来表示要创建的 Loopback 接口编号，取值范围为 0～1 023 的整数，但对于 5700-LI 系列交换机，取值范围为 0～15 的整数
3	**ip address** *ip-address* { *mask* | *mask-length* } [**sub**] 例如：[HUAWEI-Loopback1]**ip address** 192.168.0.10 255.255.255.0	为 Loopback 接口配置 IP 地址。这里的参数说明与上节介绍的以太网子接口 IP 地址的配置是一样的，参见表 4-10 中的 3 步
4	**ip verify source-address** 例如：[HUAWEI-Loopback1]**ip verify source-address**	使能 Loopback 接口对接收到的报文进行源地址合法性检查，非法源地址的报文将被丢弃。如下几种 IP 地址均为非法源地址。 （1）全 0 或全 1 的地址 （2）组播地址（D 类地址） （3）E 类地址 （4）非本机产生的环回地址（形式为 127.x.x.x） （5）A、B、C 类广播地址 （6）与入接口地址在同一网段的子网广播地址 缺省情况下，接口不对接收的报文进行源地址合法性检查，可用 **undo ip verify source-address** 命令去使能接口的该功能

（续表）

可用display interface loopback [loopback-number] 命令查看全部（当不指定 loopback-number可选参数时）或指定Loopback接口的状态信息。

【示例 】查看指定的Loopback 6接口的状态。输出信息字段说明如表 4-13所示。

```
<HUAWEI>display interface loopback 6
LoopBack6 current state : UP
Line protocol current state :UP (spoofing)
Description:HUAWEI, HUAWEI Series, LoopBack6 Interface
Route Port,The Maximum Transmit Unit is 1500
Internet Address is 1.1.1.9/32
  Input bandwidth utilization : 0.00%
  Output bandwidth utilization : 0.00%
```

表4-13 display interface loopback命令输出信息字段说明

字段	说明
LoopBack6 current state	显示对应 LoopBack 接口当前的物理状态，LoopBack 接口创建成功后物理状态一直是 Up 的
Line protocol current state	显示对应 Loopback 接口的链路协议状态。Loopback 接口创建成功后链路协议状态也一直是 Up 的
Description	显示对应 Loopback 接口描述，可以使用 Loopback 接口视图下的 **description** 命令设置
Route Port,The Maximum Transmit Unit is 1500	显示对应 Loopback 接口的最大传输单元（MTU），缺省值是 1500 字节。长度大于 MTU 的报文，将被分片后再发送。如果设置了不准分片，该报文会被丢弃
Internet Address is	显示对应 Loopback 接口的 IP 地址

4.4.3 配置NULL接口

NULL接口由系统自动创建，且只有一个编号为**0**的**NULL**接口，一直保持**Up**状态，不能进行像**IP**地址或封装其他协议那样的配置。不能用来转发报文，任何发送到该接口的网络数据报文都会被丢弃 。如果在静态路由中指定到达某一网段的下一跳为NULL0 接口，则任何发送到该网段的数据报文都会被丢弃。我们正好可以利用 NULL接口的这一特性，将需要过滤掉的报文直接发送到 NULL0 接口，而不必配置访问控制列表，更为简单。

例如：使用如下的静态路由配置命令丢弃所有去往网段192.101.0.0的报文：

```
[HUAWEI] ip route-static 192.101.0.0 255.255.0.0 NULL 0
```

4.5 以太网链路聚合

链路聚合（Link Aggregation）在华为S系列交换机中称之为Eth-Trunk，是将一组相同类型的物理以太网接口捆绑在一起的逻辑接口（就是 Eth-Trunk 接口），是用来增加带宽的一种方法。但 Eth-Trunk 口与物理以太网接口一样，也可以配置成 **Access**、**Hybrid**、**Trunk** 或 **Tunnel** 端口类型，指导它加入一个或多个 **VLAN** 中（具体参见本书第**6**章） 。

系列交换机支持手工和静态LACP两种链路聚合模式，可将两个或两个以上物理接口捆绑成一个 Eth-Trunk 接口。当聚合链路中一条链路发生故障时，故障链路上的流量还会自动分担到其他链路上，从而保证了业务传输不被中断。本节要介绍华为S系列交换机中各种链路聚合方式的配置与管理方法。

4.5.1 链路聚合特性及产品支持

随着网络规模不断扩大，用户对骨干链路的带宽和可靠性提出越来越高的要求。在传统技术中，常用

更换高速率的接口板或更换支持高速率接口板的设备的方式来增加带宽，但这种方案需要付出高额的费用，而且不够灵活。采用链路聚合技术可以在不进行硬件升级的条件下，通过将多个物理接口捆绑为一个逻辑接口，实现增加链路带宽的目的。而且，链路聚合的备份机制在有效提高可靠性的同时，还可以实现流量在不同物理链路上的负载分担。

如图4-3所示，DeviceA与DeviceB之间通过三条以太网物理链路相连，将这三条链路捆绑在一起就成为了一条逻辑链路Eth-trunk，这条逻辑链路的带宽等于原先三条以太网物理链路的带宽总和，从而达到了增加链路带宽的目的；同时，这三条以太网物理链路相互备份，有效地提高了链路的可靠性。



图4-3 链路聚合示意图

目前华为 S 系列交换机上支持手工负载分担 Eth-Trunk 链路和 LACP（Link Aggregation Control Protocol，链路聚合控制协议）Eth-Trunk链路两种聚合模式。在CSS 集群场景中支持 Eth-Trunk 接口本地流量优先转发，还支持跨设备的链路聚合E-Trunk。

1．手工负载分担模式链路聚合

手工负载分担模式是一种最基本的链路聚合方式。在该模式下，Eth-Trunk接口的建立、成员接口的加入，以及哪些接口作为活动接口完全由手工来完成的，没有链路聚合控制协议（LACP）的参与。该模式下所有活动链路都参与数据的转发，平均分担流量， 因此称为负载分担模式。如果某条活动链路故障，链路聚合组自动在剩余的活动链路中平均分担流量。

手工负载分担模式通常用于对端设备不支持LACP协议的情况下，所有S系列交换机均支持 。

2．LACP模式链路聚合

LACP模式也称“静态LACP模式”，是一种利用LACP协议进行聚合参数协商、确定活动接口和非活动接口的高级链路聚合方式。在 LACP 协议中，链路的两端分别称为Actor和Partner，双方通过LACPDU报文交互，向对端通告自己的系统优先级、系统MAC、端口优先级、端口号和操作key，对端收到LACPDU报文后将这些信息与其他端口所保存的信息进行比较，以选择能够汇聚的端口。所有 S 系列交换机均支持该特性。

在该模式下，虽然 Eth-Trunk 接口的建立，成员接口的加入也是由手工配置完成的。但与手工负载分担模式链路聚合不同的是，该模式下活动接口的选择由LACP协议报文负责。也就是说，当把一组接口加入Eth-Trunk 接口后，这些成员接口中哪些接口作为活动接口，哪些接口作为非活动接口还需要经过LACP协议报文的协商确定。

LACP模式也称为M：N模式，因为这种方式同时可以实现链路负载分担和链路冗余备份的双重功能。在链路聚合组中M条链路处于活动状态，这些链路负责转发数据并进行负载分担，另外N条链路处于非活动状态作为备份链路，不转发数据；当M条链路中有链路出现故障时，系统会从N条备份链路中选择优先级最高的接替出现故障的链路，并开始转发数据，如图4-4所示。

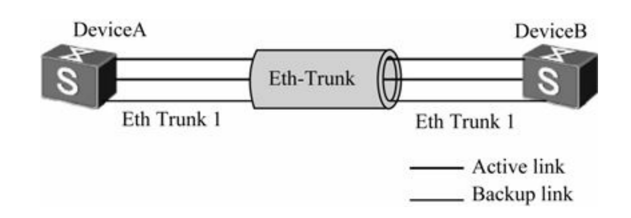


图4-4 LACP模式链路聚合示例

由上分析可知，LACP 模式与手工负载分担模式的主要区别为：LACP 模式有非活动的备份链路，而手工负载分担模式所有成员接口均处于转发状态，分担负载流量。在LACP模式下，聚合组两端的设备中LACP优先级较高的一端为主动端，LACP优先级较低的一端为被动端。区分主动端与被动端的目的是为了保证两端设备最终确定的活动接口一致，如果两端都按照本端各自的接口优先级来选择活动接口，两端所确定的活动接口很可能不一致，活动链路也就无法建立。因此首先确定主动端，被动端按照主动端侧的接口优先级来选择活动接口。

那么如何确定聚合链路的主动端和被动端呢？具体原则如下。

（1）根据聚合链路两端设备的系统LACP优先级来确定：display eth-trunk命令（具体在本章后面介绍）中“System Priority”字段和 Partner中的“SysPri”字段分别代表本端和对端设备的系统LACP优先级，值越小优先级越高，缺省情况下该值都为32 768；

（2）如果聚合链路两端的系统LACP优先级相同，则按照链路两端设备的系统MAC地址来确定，MAC地址越小优先级越高。display eth-trunk命令中“System ID”字段和Partner中的“SystemID”字段分别代表本端和对端设备的系统MAC地址。

3．堆叠场景中跨设备Eth-Trunk接口支持本地流量优先转发

交换机堆叠可增加交换机整体的转发性能，而跨交换机的 Eth-Trunk 接口可实现交换机间的备份、提高可靠性。但是由于Eth-Trunk接口通过HASH算法选择转发出接口，在交换机堆叠没有任何故障的情况下，从本交换机进入的流量很可能跨交换机进行转发。这样就增加了堆叠交换机之间的带宽承载压力，也降低了流量转发效率。此时可通过使能Eth-Trunk接口本地流量优先转发功能解决此问题。

如图4-5所示，DeviceB和DeviceC组成堆叠，堆叠交换机和DeviceA之间用Eth-Trunk连接，通过在堆叠交换机上使能Eth-Trunk接口本地流量优先转发功能可实现：

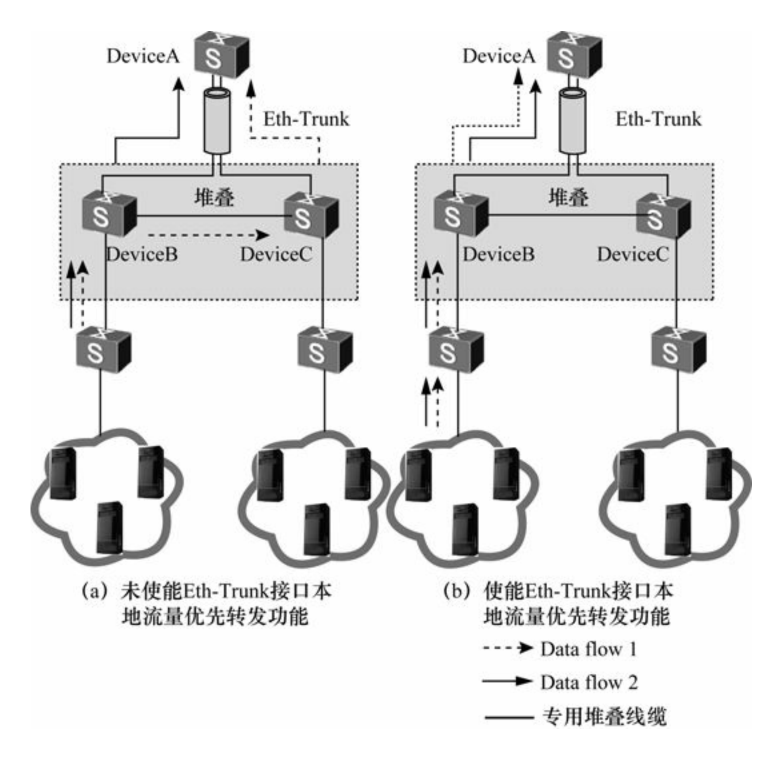


图4-5 Eth-Trunk接口本地流量优先转发示例

（1）入本设备流量从本设备转发。当Eth-Trunk接口在DeviceB有出接口且无故障时，DeviceB的Eth-Trunk接口转发表中将只包含DeviceB的出接口。这样DeviceB到DeviceA 的流量在通过 HASH 算法选择出接口时只能选中 DeviceB 的接口，流量从DeviceB本设备转发出去。

（2）入本设备流量跨设备转发。仅**S5700SI/5700EI/7700/9300/9300E/9700**系列交换机支持此功能 。当Eth-Trunk接口在DeviceB本设备无出接口或者出接口全部故障时， DeviceB 的 Eth-Trunk 转发表中将包含 Eth-Trunk 接口中所有可转发的出接口。这样DeviceB到DeviceA的流量在通过HASH算法选择出接口时将选中DeviceC上的出接口，流量将通过DeviceC跨设备转发。

4．E-Trunk

仅 **S5700** （但 **S5700LI/5700S-LI** 除外）、**S6700/7700/9300/9300E/9700**系列支持**E-Trunk** 。E-Trunk（Enhanced Trunk，增强 Trunk）应用于CE接入网络时在CE与双PE间实现链路保护。如图 4-6 所示，CE 分别通过一条 LACP 模式的Eth-Trunk1和Eth-Trunk2与PE1、PE2相连。这两个Eth-Trunk构成一条E-Trunk，在PE1与PE2之间实现链路聚合组的备份，提高网络可靠性。

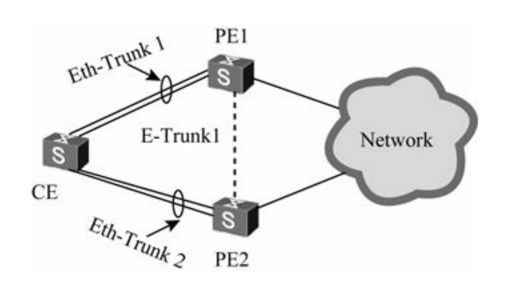


图4-6 E-Trunk组网示例

4.5.2 手工负载分担模式链路聚合配置任务

通过配置链路聚合可以达到负载分担、增加带宽、提高可靠性的目的。本节介绍手工负载分担模式链路聚合的配置与管理方法。首先介绍它的基本配置任务，也可以说是它的基本配置思路。

在整个手工负载分担模式链路聚合的配置中，可分为以下几项基本的配置任务。

1．创建链路聚合组

这是最先进行的配置任务。每个链路聚合组唯一对应着一个逻辑接口，即Eth-Trunk接口。配置链路聚合时首先要创建这样一个Eth-Trunk接口。

2．配置链路聚合模式为手工负载分担模式

根据是否启用链路聚合控制协议 LACP，链路聚合分为手工负载分担模式和 LACP模式。在本节介绍的手工负载分担模式下，Eth-Trunk的建立、成员接口的加入完全由手工来配置，而且链路聚合组中的所有活动链路都参与数据的转发，平均分担流量 。手工负载分担模式通常应用在对端设备不支持LACP协议的情况下。

注意

改变Eth-Trunk工作模式前应确保该Eth-Trunk中没有加入任何成员接口，否则无法更改Eth-Trunk的工作模式。

3．将成员接口加入聚合组

创建了链路聚合组，并且配置好了聚合模式后，就要向聚合组中添加以太网接口成员了。向聚合组中加入成员接口可基于 Eth-Trunk 接口视图配置，也可基于成员接口视图配置，用户根据需要选择以下配置之一即可。将成员接口加入 Eth-Trunk 时，需要注意以下问题。

（1）成员接口以下属性必须是缺省值 ：链路类型（Hybrid 类型）、最大广播流量百分比 、最大组播流量百分比、最大未知单播流量百分比、所属VLAN、VLAN-Mapping、VLAN-Stacking、QinQ协议号、接口优先级、是否允许BPDU报文通过、MAC地址学习功能、静态加入组播组、广播报文丢弃、未知组播报文丢弃、未知单播报文丢弃、NDP功能和NTDP功能等。

（2）每个Eth-Trunk接口下最多只可以包含8个成员以太网接口。

（3）成员以太网接口不能配置任何业务和静态**MAC**地址 ，因为此时这些成员接口上的业务和MAC地址必须在Eth-Trunk接口上配置。

（4）在成员以太网接口加入Eth-Trunk时，必须为缺省的**Hybrid**类型 。

（5）Eth-Trunk接口不能嵌套，即一个**Eth-Trunk**接口不能是另一个**Eth-Trunk**接口的成员 。

（6）一个以太网接口最多只能属于一个 **Eth-Trunk** 口中 ，如果需要加入其他Eth-Trunk接口，必须先退出原来的Eth-Trunk接口。

（7）一个**Eth-Trunk**接口中的成员接口必须是同一类型 ，如要么同时为百兆以太网接口、要么同时为千兆以太网接口等。

（8）如果本地设备创建了 **Eth-Trunk** 接口，与成员接口直连的对端接口也必须捆绑创建 **Eth-Trunk** 接口 ，否则两端不能正常通信。也就是链路聚合必须在链路两端同时配置。

（9）当成员以太网接口加入Eth-Trunk接口后，学习MAC地址时是按照Eth-Trunk接口来进行的，而不是按照各成员接口来学习的。

（10）**Eth-Trunk**链路两端相连的各成员以太网接口的数量、速率、双工模式、超大帧支持、流量控制等属性配置必须一致 。

4．（可选）配置活动接口数阈值

本项配置任务仅为保证Eth-Trunk接口的状态和带宽，以减小个别成员链路的状态变化对整条聚合链路带来的影响。设置活动接口数下限阈值是为了保证最小带宽，当前活动链路数目小于下限阈值时，Eth-

Trunk 接口的状态转为 Down。这时其中的成员不再形成聚合状态，而是恢复各处独立物理接口状态。但活动接口数上限阈值不适用于手工负载分担模式 。

5.（可选）配置负载分担方式

缺省情况下，Eth-Trunk 的负载分担是按流进行的，以保证包的正确顺序，即保证了同一数据流的帧在同一条物理链路转发，而不同数据流在不同的物理链路上转发从而实现分担负载。华为S系列交换机都可以配置普通负载分担模式，即基于报文的源/目的IP地址或源/目的MAC地址来分担负载；但对于S5710EI/5700HI/6700/7700/9300/ 9300E/9700系列还可针对二层报文、IP 报文和 MPLS 报文配置增强型的负载分担模式。由于负载分担只对出方向的流量有效，因此链路两端接口的负载分担方式可以不一致，互不影响 。由于增强型的负载分担模式配置比较复杂，且在一般的企业中应用比较少，故在此不作介绍。

目前华为S系列交换机所支持的普通负载分担方式如下。

（1）dst-ip（目的IP地址）：从报文中的目的IP地址、出端口的TCP/UDP端口号中分别选择指定位的3位数值进行异或运算，根据运算结果选择Eth-Trunk 链路表中对应的出接口。结果是，来自同一个源 **IP** 地址而要发送到不同目的 **IP** 地址的数据包将在**Eth-Trunk**链路中的不同端口上发送 ，以此来实现负载均衡。但是来自不同源**IP**地址但相同目的**IP**地址的数据包总是在**Eth-Trunk**链路的同一个端口上发送 。

（2）dst-mac（目的MAC地址）：从报文中的目的MAC地址、VLAN ID、以太网类型及入端口信息中分别选择指定位的 3 位数值进行异或运算，根据运算结果选择Eth-Trunk 链路表中对应的出接口。结果是，到达同一个 **MAC** 地址的数据包将在**Eth-Trunk**链路中的同一个端口上进行转发，不同目的**MAC**地址的数据包采用不同端口进行转发 ，以此来实现负载均衡。

（3）src-ip（源IP地址）：从报文中的源IP地址、入端口的TCP/UDP端口号中分别 选择指定位的3 位数值进行异或运算，根据运算结果选择Eth-Trunk 链路表中对应的出接口。结果是，来自不同 **IP** 地址的数据包将在 **Eth-Trunk** 链路中的不同端口上进行转发 ，以此来实现负载均衡。但是来自同一个源**IP**地址但目的**IP**地址不一样的数据包总是在**Eth-Trunk**链路的同一个端口上发送 。

（4）src-mac（源MAC地址）：从报文中的源MAC地址、VLAN ID、以太网类型及入端口信息中分别选择指定位的 3 位数值进行异或运算，根据运算结果选择 Eth-Trunk链路表中对应的出接口。结果是，来自不同**MAC**地址主机的数据包将在**Eth-Trunk**链路中的不同端口上进行转发 ，但是来自同一个**MAC**地址主机的数据包总是在**Eth-Trunk**链路中相同的端口进行转发， 以此来实现负载均衡。

（5）src-dst-ip（源IP地址与目的IP地址的异或）：从报文中的目的IP地址、源IP地址两种负载分担模式的运算结果进行异或运算，根据运算结果选择 Eth-Trunk 链路表中对应的出接口。这种转发方法是一种结合源和目的IP地址进行负载分配的转发方法。这在不清楚在特定交换机上是采用基于源IP地址转发还是采用基于目的IP地址转发更适合时可以采用。在这种基于源和目的IP地址的均衡方法中，从IP地址A到达IP地址B、从IP地址A到达IP地址C，以及从IP地址C到达IP地址B的数据包使用Eth-Trunk链路中不同的端口进行转发。

（6）src-dst-mac（源 MAC 地址与目的 MAC 地址的异或）：从报文中的目的 MAC地址、源MAC地址、VLAN ID、以太网类型及入端口信息中分别选择指定位的 3位数值进行异或运算，根据运算结果选择Eth-Trunk 链路表中对应的出接口。这种转发方法是一种结合源和目的 MAC 地址进行负载分配的转发方法。这在不清楚在特定交换机上是采用基于源 MAC 地址转发还是采用基于目的 MAC 地址转发更适合时可以采用。在这种基于源和目的MAC地址的均衡方法中，从主机A到达主机B、从主机A到达主机C，以及从主机C到达主机B的数据包可以使用Eth-Trunk链路中不同的端口进行转发。

4.5.3 手工负载分担模式链路聚合配置与管理

上节介绍的五项手工负载分担模式链路聚合的主要配置任务所对应的具体配置与管理步骤如表4-14所示。

表4-14 手工负载分担模式链路聚合的配置与管理步骤

配置任务	步骤	命令	说明
创建链路聚合组	1	**system-view** 例如：<HUAWEI> **system-view**	进入系统视图
	2	**interface eth-trunk** *trunk-id* 例如：[HUAWEI] **interface eth-trunk** 10	创建 Eth-Trunk 接口，并进入 Eth-Trunk 接口视图。参数 *trunk-id* 用来指定所创建的 Eth-Trunk 接口编号，但不同系列产品的取值范围有所不同，如 S2700EI 系列为 0～13 的整数；S2700SI 系列为 0～2 的整数，S3700 系列为 0～19 的整数，S5700SI 系列为 0～31 的整数，S5700LI/5700S-LI/5710EI/5700EI/5700HI/6700 系列为 0～63 的整数，S7700/9300/9300E/9700 系列为 0～27 的整数 可用 **undo interface eth-trunk** *trunk-id* 来删除 Eth-Trunk 接口，但在删除 Eth-Trunk 时，Eth-Trunk 接口中不能有成员以太网接口

（续表）

<table>
<tr><th>配置任务</th><th>步骤</th><th>命令</th><th colspan="2">说明</th></tr>
<tr><td>配置链路聚合模式为手工负载分担模式</td><td>3</td><td>mode manual load-balance
例如：[HUAWEI-Eth-Trunk10]mode manual load-balance</td><td colspan="2">配置 Eth-Trunk 接口的工作模式为手工负载分担模式。该模式为链路聚合组的创建和接口的加入都需要手工配置，即系统不会自动形成链路聚合，也不会自动根据某些条件加入所需的成员以太网接口。
因为缺省情况下，Eth-Trunk 接口的工作模式为手工负载分担模式，所以本项配置任务一般情况下是可以不进行的。如果当前交换机上某链路聚合组已配置成其他模式，则可用 undo mode 命令恢复对应 Eth-trunk 接口的工作模式为缺省的手工负载分担模式
【注意】配置时需要保证本端和对端的聚合模式一致。即如果本端配置为手工负载分担模式，那么对端设备也必须要配置为手工负载分担模式。另外更改 Eth-trunk 接口的工作模式需要在确保 Eth-trunk 接口中不包含任何成员以太网接口
另外，本命令为覆盖式命令，即当多次执行该命令后以最后设定的模式为最终 Eth-Trunk 接口工作模式</td></tr>
<tr><td rowspan="4">将成员接口加入聚合组（有两种方式添加，根据需要选择其一）</td><td>4</td><td>trunkport interface-type { interface-number1 [to interface-number2] } &<1-8>
例如：[HUAWEI-Eth-Trunk10]trunkport gigabitethernet 0/0/1 to 0/0/3</td><td colspan="2">在 Eth-Trunk 接口视图下添加成员以太网接口（必须为 Hybrid 类型）。接口在加入 Eth-Trunk 时，接口的部分属性必须是缺省值，否则将无法加入。命令中的参数和选项说明如下。
（1）interface-type：指定要加入的成员以太网接口的接口类型
（2）interface-number1：指定要加入的成员以太网接口的第一个接口的编号
（3）interface-number2：可选参数，指定要加入的成员以太网接口的最后一个接口的编号
（4）&<1-8>：表示前面的{ interface-number1 [to interface-number2] }参数最多可有 8 个，因为每个 Eth-Trunk 接口下最多可以加入 8 个成员接口。但不同类型的接口不能加入同一个 Eth-Trunk 接口中
缺省情况下，Eth-Trunk 接口没有加入任何成员接口，可用 undo trunkport interface-type { interface-number1 [to interface-number2] } &<1-8>命令在 Eth-Trunk 接口视图下删除指定的成员接口</td></tr>
<tr><td rowspan="3">4</td><td>quit
例如：[HUAWEI-Eth-Trunk10] quit</td><td rowspan="3">在接口视图下添加成员以太网接口</td><td>退出 Eth-Trunk 接口视图，返回系统视图</td></tr>
<tr><td>interface interface-type interface-number
例如：[HUAWEI] interface GigabitEthernet0/0/1</td><td>键入要加入 Eth-Trunk 接口的 Hybrid 类型成员以太网接口，进入接口视图</td></tr>
<tr><td>eth-trunk trunk-id
例如：[HUAWEI-GigabitEthernet0/0/1] eth-trunk 10</td><td>将当前接口加入指定的 Eth-Trunk 接口中。接口在加入 Eth-Trunk 时，接口的部分属性必须是缺省值
缺省情况下，当前接口不属于任何 Eth-Trunk，可用 undo eth-trunk 命令将当前接口从指定 Eth-Trunk 中删除</td></tr>
</table>

（续表）

配置任务	步骤	命令	说明
（可选）配置活动接口数阈值	5	**interface eth-trunk** *trunk-id* 例如：[HUAWEI] **interface eth-trunk** 10	（可选）进入Eth-Trunk接口视图，**如果前面是在Eth-Trunk接口视图下添加成员接口的，则不要进行此步**
	6	**least active-linknumber** *link-number* 例如：[HUAWEI-Eth-Trunk10] **least active-linknumber** 4	在Eth-Trunk接口视图下配置链路聚合**活动接口数**下限阈值。参数*link-number*用来指定链路聚合活动接口数下限阈值，除S2700SI系列的取值范围为1～4的整数外，其他支持VRP系统的Sx700系列的取值范围均为1～8的整数 **本端和对端设备的活动接口数下限阈值可以不同**。如果下限阈值不同，以下限阈值数值较大的一端为准。执行本命令后，**当活动链路数低于所配置的下限阈值时，Eth-Trunk接口状态变为Down，所有的Eth-Trunk接口成员不再转发数据**，能够避免因活动链路数目减少负载过大而出现丢包的现象；当Eth-Trunk接口中活动接口数达到设置的下限阈值时，Eth-Trunk接口状态将变为Up **本命令为覆盖式命令**，当多次配置活动接口数下限阈值后，以最后一次配置为最终下限阈值 缺省情况下，活动接口数下限阈值为1，可用**undo least active-linknumber**命令恢复聚合组活动接口数目的下限阈值为缺省值
（可选）配置普通负载分担方式	7	**load-balance** { **dst-ip** | **dst-mac** | **src-ip** | **src-mac** | **src-dst-ip** | **src-dst-mac** } 例如：[HUAWEI-Eth-Trunk10]**load-balance src-ip**	配置Eth-Trunk接口的普通负载分担方式。命令中的选项说明如下（有关各种负载分担模式的说明参见4.5.2节）： （1）**dst-ip**（目的IP地址）：多选一选项，根据报文中的目的IP地址进行负载分担 （2）**dst-mac**（目的MAC地址）：多选一选项，根据报文中的目的MAC地址进行负载分担 （3）**src-ip**（源IP地址）：多选一选项，根据报文中的源IP地址进行负载分担 （4）**src-mac**（源MAC地址）：多选一选项，根据报文中的源MAC地址进行负载分担 （5）**src-dst-ip**（源IP地址与目的IP地址）：多选一选项，同时根据报文中的源IP地址与目的IP地址进行负载分担 （6）**src-dst-mac**（源MAC地址与目的MAC地址）：多选一选项，同时根据报文中的源MAC地址与目的MAC地址进行负载分担 缺省情况下，Eth-Trunk接口的负载分担模式为**src-dst-ip**，可用**undo load-balance**命令恢复Eth-Trunk接口的负载分担模式为缺省的**src-dst-ip**模式

可用display eth-trunk [trunk-id [interface interface-type interface-number |verbose]] 命令查看所有或者指定Eth-Trunk接口的摘要或者详细（选择verbose可选项时）配置信息；可用display trunkmembership eth-trunk trunk-id命令查看指定Eth-Trunk接口的成员接口信息。

【示例 1】查看接口Eth-Trunk 1手工负载分担模式的配置信息。输出信息字段说明如表4-15所示。

```
<HUAWEI>display eth-trunk 1
Eth-Trunk1's state information is:
  WorkingMode: NORMAL    Hash Arithmetic: According to SA-XOR-DA
  Least Active-linknumber:2    Max Bandwidth-affected-linknumber:8
  Operate Status: up    Number Of Up Ports In Trunk:2
--------------------------------------------------------------------------------
  PortName    Status    Weight
  GigabitEthernet1/0/1    Up    1
  GigabitEthernet1/0/2    Up    1
```

表4-15 display eth-trunk命令输出信息字段说明

字段	说明
WorkingMode	显示对应 Eth-Trunk 接口的工作模式：NORMAL 表示手工负载分担模式，STATIC 表示 LACP 模式
Hash arithmetic	显示 Eth-Trunk 接口的 hash 算法。由 **load-balance** 命令配置的接口负载分担模式决定
Least active-linknumber	显示对应 Eth-Trunk 接口处于 Up 状态的成员链路的下限阈值
Max bandwidth-affected-linknumber	显示影响对应 Eth-Trunk 接口带宽的最大连接数
Operate status	显示对应 Eth-Trunk 接口的状态：Up 表示接口处于正常启动的状态，Down 表示接口在物理上出现故障
Number Of Up Ports in Trunk	显示对应 Eth-Trunk 接口中处于 Up 状态的成员接口数
PortName	显示成员接口名
Status	显示手工负载分担模式下，本地成员接口的状态：Up 表示对应成员接口处于正常启动的状态，Down 表示对应成员接口在物理上出现故障
Weight	显示对应成员接口的权重

【示例 2】查看Eth-Trunk 2接口的成员接口信息。输出信息字段说明如表 4-16所示。

<HUAWEI>display trunkmembership eth-trunk 2

Trunk ID: 2

used status: VALID

TYPE: ethernet

Working Mode : Normal

Number Of Ports in Trunk = 2

Number Of UP Ports in Trunk = 2

operate status: up

Interface GigabitEthernet1/0/1, valid, operate up, weight=1

Interface GigabitEthernet1/0/2, valid, operate up, weight=1

表4-16 display trunkmembership eth-trunk命令的输出信息字段说明

字段	说明
Trunk ID	显示对应 Eth-Trunk 接口的编号
used status	显示对应 Eth-Trunk 接口的状态：VALID 表示对应 Eth-Trunk 接口有效，INVALID 表示对应 Eth-Trunk 接口无效
TYPE	显示对应 Eth-Trunk 接口的接口类型
Working Mode	显示对应 Eth-Trunk 接口的负载分担模式：Normal 表示普通负载分担模式，STATIC 表示静态 LACP 负载分担模式
Number Of Ports in Trunk	显示对应 Eth-Trunk 接口中包含的成员接口个数
Number Of UP Ports in Trunk	显示对应 Eth-Trunk 接口中包含的处于开启状态的接口个数
operate status	显示成员接口的状态：Down 表示关闭成员接口，Up 表示开启成员接口
Interface GigabitEthernet1/0/1, valid,operate up,weight=1	显示成员接口 GigabitEthernet1/0/1 的状态，包括以下两种。 （1）有效状态：valid 表示成员接口有效，invalid 表示成员接口无效 （2）操作状态：operate down 表示关闭成员接口，operate up 表示开启成员接口

4.5.4 手工负载分担模式链路聚合配置示例

本示例拓扑结构如图 4-7 所示，SwitchA 和 SwitchB 通过以太网链路分别连接VLAN10和VLAN20，且SwitchA和SwitchB之间有较大的数据流量。现希望SwitchA和SwitchB之间能够提供较大的链路带宽使相同VLAN间互相通信。同时用户也希望能够提供一定的冗余度，保证数据传输和链路的可靠性。

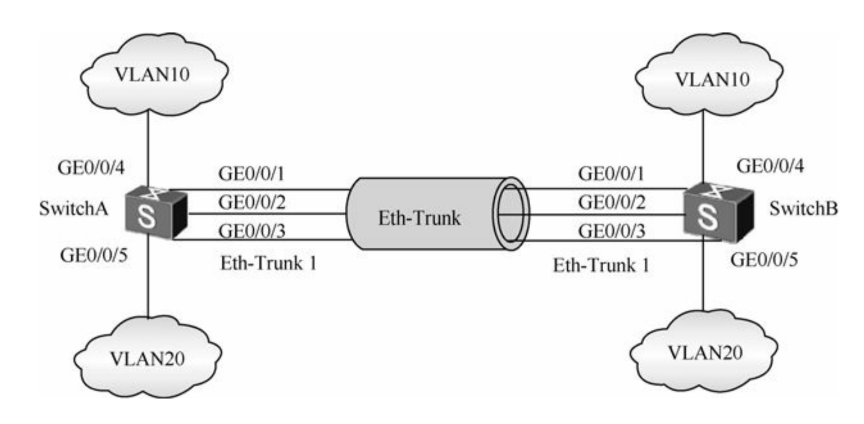


图4-7 手工负载分担模式链路聚合配置示例拓扑结构

因为本示例并没有要求提供链路备份功能，所以可以采用相对简单的手工负载分担链路聚合方式来进行配置。五项配置任务在表4-14中已有详细介绍，不再赘述。但要注意的是，3个成员接口GE0/0/1～0/0/3在加入Eth-Trunk接口前一定要恢复为缺省配置（特别是为缺省的Hybrid类型）。另外，最好将这些成员接口从缺省的VLAN1退出或关闭，避免出现广播风暴。因为本示例中SwitchA与SwitchB的配置是对称的，所以下面仅以SwitchA为例介绍具体的配置步骤。

（1）在SwitchA上创建Eth-Trunk接口（此处为Eth-Trunk1），指出为手工负载分担模式，并加入成员接口GE0/0/1～0/0/3。

```
<HUAWEI>system-view
[HUAWEI] sysname SwitchA
[SwitchA] interface Eth-Trunk1
[SwitchA-Eth-Trunk1] modemanual load-balance
[SwitchA-Eth-Trunk1] trunkport gigabitethernet 0/0/1 to 0/0/3
[SwitchA-Eth-Trunk1] quit
```

（2）创建VLAN并将其他端口加入VLAN 10或VLAN 20中。有关VLAN的具体创建和端口加入方法参见本书第6章相关内容。

```
[SwitchA] vlan batch 10 20    #---批量创建VLAN 10和VLAN 20
[SwitchA] interface gigabitethernet 0/0/4
[SwitchA-GigabitEthernet0/0/4] port link-type trunk    #---设置GE0/0/4接口类型为Trunk类型
[SwitchA-GigabitEthernet0/0/4] port trunk allow-pass vlan 10    #---允许VLAN 10的报文通过
[SwitchA-GigabitEthernet0/0/4] quit
[SwitchA] interfacegigabitethernet 0/0/5
[SwitchA-GigabitEthernet0/0/5] port link-type trunk
[SwitchA-GigabitEthernet0/0/5] port trunk allow-pass vlan 20
[SwitchA-GigabitEthernet0/0/5] quit
```

（3）配置Eth-Trunk1接口为Trunk类型，并允许VLAN10和VLAN20通过。

说明

Eth-Trunk接口与物理以太网接口一样，也可以根据实际需要配置成各种端口类型（因为是设备之间的链路，所以通常是Trunk或者带标签的Hybrid类型），允许来自一个或多个VLAN的数据通过。具体配置方法参见本书第6章。

[SwitchA] interface Eth-Trunk 1

[SwitchA-Eth-Trunk1] port link-type trunk

[SwitchA-Eth-Trunk1] port trunk allow-pass vlan 10 20

（4）配置Eth-Trunk1的负载分担方式为src-dst-mac，即基于报文中的源MAC地址和目的MAC地址方式，因为这里是二层VLAN报文。

[SwitchA-Eth-Trunk1] load-balance src-dst-mac

[SwitchA-Eth-Trunk1] quit

配置好后，可在任意视图下执行display eth-trunk1命令检查Eth-Trunk是否创建成功及成员接口是否正确加入。本示例执行此命令后的输出信息如下（注意输出信息中的粗体字部分）：

[SwitchA] display eth-trunk 1

Eth-Trunk1's state information is:

WorkingMode:NORMAL   Hash arithmetic: According to SA-XOR-DA

Least Active-linknumber: 1   Max Bandwidth-affected-linknumber: 8

Operate status: up   Number Of Up Port In Trunk: 3

PortName   Status   Weight

GigabitEthernet0/0/1   Up   1

GigabitEthernet0/0/2   Up   1

GigabitEthernet0/0/3   Up   1

从以上信息看出Eth-Trunk 1中包含3个成员接口GigabitEthernet0/0/1、GigabitEthernet0/0/2和GigabitEthernet0/0/3。成员接口的状态都为Up，表明配置是成功的。

4.5.5 LACP模式链路聚合配置任务

LACP 模式链路聚合与上节介绍的手工负载分担模式链路聚合相比，最大的优势就是既可以实现负载分担，又可以同时实现链路备份。下面先介绍它的主要配置任务。

在整个LACP模式链路聚合的配置中，可分为以下几项基本的配置任务。

1．创建链路聚合组

这一步与上节介绍的手工负载分担模式链路聚合配置中的第一项配置任务一样。每个链路聚合组唯一对应着一个逻辑接口，即Eth-Trunk接口。配置LACP模式链路聚合时也首先要创建这样一个Eth-Trunk接口。

2．配置链路聚合模式为LACP模式

在LACP模式下，同样需手工创建Eth-Trunk，手工加入Eth-Trunk成员接口，但活动接口的选择是由LACP协商确定的，配置相对灵活。改变**Eth-Trunk**工作模式前应确保该**Eth-Trunk**中没有加入任何成员接口，否则无法更改Eth-Trunk的工作模式。

3．将成员接口加入聚合组

向聚合组中加入成员接口可基于Eth-Trunk接口视图配置，也可基于成员接口视图配置，根据需要选择其一即可。在添加成员接口时同样要注意上节第（3）项配置任务中的注意事项。

4．（可选）配置活动接口数阈值

为保证 Eth-Trunk 接口的状态和带宽，可以设置活动接口数的阈值，以减小成员链 路的状态变化带来的影响。在LACP模式的聚合链路中可以设置以下两个阈值。

（1）活动接口数下限阈值：设置活动接口数下限阈值是为了保证最小带宽，当前活动链路数目小于下限阈值时，Eth-Trunk接口的状态转为Down。

（2）活动接口数上限阈值：设置活动接口数上限阈值的目的是在保证带宽的情况下提高网络的可靠性（这在上节介绍的手工负载分担模式中是没有的配置，因在手工负载分担模式中，各链路都是用来进行负载分担的，没有备份链路）。当前活动链路数目达到上限阈值时，再向Eth-Trunk中添加成员接口，不会增加Eth-Trunk活动接口的数目，超过上限阈值的链路状态将被置为Down。

5. （可选）配置负载分担方式

缺省情况下，Eth-Trunk的负载分担方式是同一数据流的帧在同一条物理链路转发，不同数据流的帧在不同的物理链路上转发，保证了数据包传递的正确顺序。也可以配置普通负载分担模式，基于报文的IP地址或MAC地址来分担负载；对于二层报文、IP报文和MPLS报文还可以配置增强型的负载分担模式。由于负载分担只对出方向的流量有效，因此，链路两端接口的负载分担模式可以不一致，两端互不影响。

6. （可选）配置系统LACP优先级

系统 LACP 优先级是为了区分链路聚合两端设备优先级的高低而配置的参数。在LACP模式下，两端设备所选择的活动接口必须保持一致，否则链路聚合组就无法建立。而要想使两端活动接口保持一致，可以使其中一端具有更高的优先级，另一端根据高优先级的一端来选择活动接口即可。

7. （可选）配置接口LACP优先级

LACP模式下可以通过配置接口LACP优先级来区分不同接口被选为活动接口的优先程度，优先级高的接口将优先被选为活动接口。

8. （可选）配置LACP抢占

在LACP模式下，当活动链路中出现故障链路时系统会从备用链路中选择优先级最高的链路替代故障链路；如果被替代的故障链路恢复了正常，而且该链路的优先级又高于替代自己的链路。这时如果使能了LACP优先级抢占功能，高优先级链路会抢占低优先级链路，回切到活动状态，否则，系统不会重新选择活动接口，故障恢复后的链路将作为备用链路。在进行优先级抢占时，系统将根据主动端接口的优先级进行抢占。

在这里还涉及一个概念——抢占延时，也就是抢占等待时间，是指在LACP模式的Eth-Trunk中非活动接口切换为活动接口需要等待的时间。配置抢占延时可以避免由于某些链路状态频繁变化而导致Eth-Trunk数据传输不稳定的情况。

9. （可选）配置接收LACP报文超时时间

如果对端链路聚合组的某个成员端口发生自环或其他故障，而本端 Eth-Trunk 接口不能及时感知对端成员口状态的变化，就会导致本端转发数据时仍按照本端链路组中活动接口进行负载分担，造成发生故障链路上数据流量的丢失。配置接口接收LACP报文的超时时间后，如果本端成员口在设置的超时时间内未收到对端发送的LACP协议报文，则认为对端不可达，本端成员口状态立即变为Down，不再转发数据。

与LACP模式链路聚合相关参数的缺省配置如表4-17所示。

表4-17 链路聚合参数缺省值

参数	缺省值
链路聚合模式	手工负载分担模式
活动接口数上限阈值	8
活动接口数下限阈值	1
系统 LACP 优先级	32768
接口 LACP 优先级	32768
LACP 抢占	Disabled
LACP 抢占等待时间	30s
接收 LACP 报文超时时间	90s
Eth-Trunk 接口本地流量优先转发	Enabled

4.5.6 LACP模式链路聚合配置与管理

上节介绍的九项 LACP 模式链路聚合配置任务所对应的具体配置与管理步骤如表4-18所示。

表4-18 LACP模式链路聚合的配置与管理步骤

<table>
<tr><th>配置任务</th><th>步骤</th><th>命令</th><th colspan="2">说明</th></tr>
<tr><td rowspan="2">创建链路聚合组</td><td>1</td><td>system-view
例如：<HUAWEI> system-view</td><td colspan="2">进入系统视图</td></tr>
<tr><td>2</td><td>interface eth-trunk trunk-id
例如：[HUAWEI] interface eth-trunk 10</td><td colspan="2">创建 Eth-Trunk 接口，并进入 Eth-Trunk 接口视图，如果该 Eth-Trunk 已经存在，本命令用来进入 Eth-Trunk 接口视图。其他说明参见 4.5.3 节表 4-14 中的第 2 步</td></tr>
<tr><td>配置 LACP 聚合模式</td><td>3</td><td>mode lacp
例如：[HUAWEI-Eth-Trunk10] mode lacp</td><td colspan="2">配置 Eth-Trunk 接口的工作模式为 LACP 模式。缺省情况下，Eth-Trunk 的工作模式为手工负载分担模式。其他说明参见 4.5.3 节表 4-14 中的第 3 步</td></tr>
<tr><td rowspan="4">将成员接口加入聚合组（有两种方式添加，根据需要选择其一）</td><td rowspan="4">4</td><td>trunkport interface-type { interface-number1 [to interface-number2] } &<1-8>
例如：[HUAWEI-Eth-Trunk10]trunkport gigabitethernet 0/0/1 to 0/0/3</td><td colspan="2">在 Eth-Trunk 接口视图下添加成员以太网接口（Hybrid 类型）。接口在加入 Eth-Trunk 时，接口的部分属性必须是缺省值，否则将无法加入。其他说明参见 4.5.3 节表 4-14 中的第 4 步</td></tr>
<tr><td>quit
例如：[HUAWEI-Eth-Trunk10] quit</td><td rowspan="3">在成员接口视图下添加成员以太网接口</td><td>退出 Eth-Trunk 接口视图，返回系统视图</td></tr>
<tr><td>interface interface-type interface-number
例如：[HUAWEI] interface GigabitEthernet0/0/1</td><td>键入要加入 Eth-Trunk 接口的 Hybrid 类型成员以太网接口，进入接口视图</td></tr>
<tr><td>eth-trunk trunk-id
例如：[HUAWEI-GigabitEthernet0/0/1] eth-trunk 10</td><td>将当前接口加入指定的 Eth-Trunk 接口中。接口在加入 Eth-Trunk 时，接口的部分属性必须是缺省值，否则将无法加入
缺省情况下，当前接口不属于任何 Eth-Trunk，可用 undo eth-trunk 命令将当前接口从指定 Eth-Trunk 中删除</td></tr>
</table>

（续表）

配置任务	步骤	命令	说明
(可选)配置活动接口数阈值	5	**interface eth-trunk** *trunk-id* 例如：[HUAWEI] **interface eth-trunk** 10	（可选）进入 Eth-Trunk 接口视图，**如果前面是在 Eth-Trunk 接口视图下添加成员接口的，则不要进行此步**
	6	**least active-linknumber** *link-number* 例如：[HUAWEI-Eth-Trunk10] **least active-linknumber** 4	在 Eth-Trunk 接口视图下配置链路聚合活动接口数下限阈值。其他说明参见 4.5.3 节表 4-14 中的第 6 步
	7	**max active-linknumber** *link-number* 例如：[HUAWEI-Eth-Trunk10] **max active-linknumber** 6	配置链路聚合活动接口数上限阈值，取值范围为 1～8 的整数 【说明】配置此命令后，如果当前活动接口数已经达到配置的上限阈值，再新增加成员接口不会影响当前的活动接口数目。当加入到汇聚端口中的成员数目小于指定的活动接口的最大数目的时候，没有备份端口。本端和对端设备的活动接口数上限阈值可以不同。如果上限阈值不同，以上限阈值数值较小的一端为准 **本命令为覆盖式命令**，当多次配置链路聚合组活动接口数目上限阈值后，以最后一次配置为最终上限阈值。剩余的链路作为备份链路缺省情况下，活动接口数上限阈值为 8，可用 **undo max active-linknumber** 命令恢复聚合组活动接口数目的上限阈值为缺省值
(可选)配置普通负载分担方式	8	**load-balance** { **dst-ip** | **dst-mac** | **src-ip** | **src-mac** | **src-dst-ip** | **src-dst-mac** } 例如：[HUAWEI-Eth-Trunk10]**load-balance src-ip**	配置 Eth-Trunk 接口的普通负载分担方式。其他说明参见 4.5.3 节表 4-14 中的第 7 步。有关各种负载分担模式说明参见 4.5.2 节相关内容
(可选)配置系统 LACP 优先级	9	**lacp priority** *priority* 例如：[HUAWEI-Eth-Trunk10] **lacp priority** 10	配置当前设备的系统 LACP 优先级，取值范围为 0～65 535 的整数，**值越小优先级越高**。在两端设备中选择系统 LACP 优先级较小一端作为主动端，如果系统 LACP 优先级相同则选择 MAC 地址较小的一端作为主动端 【说明】配置系统优先级是为了区别本端设备与对端设备优先级的高低，系统优先级高的将被选作链路聚合组的主动端，即按照主动端设备的链路接口来选择活动接口 缺省情况下，系统 LACP 优先级为 32768，可用 **undo lacp priority** 命令恢复本端设备的系统 LACP 优先级值为缺省值
(可选)配置接口 LACP 优先级	10	**quit** 例如：[HUAWEI-Eth-Trunk10] **quit**	退出 Eth-Trunk 接口视图，返回系统视图
	11	**interface** *interface-type interface-number* 例如：[HUAWEI] **interface** gigabitethernet0/0/1	键入要配置接口 LACP 优先级的成员接口，进入接口视图

（续表）

配置任务	步骤	命令	说明
（可选）配置接口LACP优先级	12	**lacp priority** *priority* 例如：[HUAWEI-GigabitEthernet0/0/1] **lacp priority** 10	配置当前成员接口的LACP优先级，取值范围为0～65 535的整数，**值越小优先级越高**，优先级高的将被选作活动接口 【说明】如果在LACP模式下执行了表中第5步的**max actvie-linknumber**命令配置了活动接口数目上限阈值，当手工加入链路组的接口数超过了该阈值的限制，就需要选择哪些接口为活动接口，而设置接口LACP优先级即可保证在LACP模式下高优先级的接口成为活动接口。如果没有执行表中第5步中的**max actvie-linknumber**命令，活动接口上限阈值为最大值8。只要手工加入的成员接口数值小于8，就不需要选择活动接口，所有的接口都处于活动状态 缺省情况下，接口LACP优先级为32768，可用**undo lacp priority**命令恢复本接口的LACP优先级值为缺省值
（可选）配置LACP抢占	13	**quit** 例如：例如：[HUAWEI-GigabitEthernet0/0/1] **quit**	退出接口视图，返回系统视图
	14	**interface eth-trunk** *trunk-id* 例如：[HUAWEI] **interface eth-trunk** 10	进入Eth-Trunk接口视图
	15	**lacp preempt enable** 例如：[HUAWEI-Eth-Trunk10] **lacp preempt enable**	使能当前Eth-Trunk接口的LACP抢占功能。在进行优先级抢占时，系统将根据主动端接口的优先级进行抢占。但要求Eth-Trunk两端LACP抢占功能使能情况配置一致，即统一使能或不使能缺省情况下，LACP抢占处于去使能状态 【说明】在LACP静态模式下，如果对应Eth-Trunk接口使能了抢占功能，则当活动链路中出现故障链路时，系统会从备用链路中选择优先级最高的链路替代故障链路；如果被替代的故障链路恢复了正常，而且该链路的优先级又高于替代自己的链路时，高优先级链路会抢占低优先级链路，切回到活动状态。如果不使能优先级抢占功能，系统不会重新选择活动接口，故障恢复后的链路将作为备用链路 缺省情况下，优先级抢占处于禁止状态，可用 **undo lacp preempt enable**命令禁止LACP模式下的LACP优先级抢占功能
	16	**lacp preempt delay** *delay-time* 例如：[HUAWEI-Eth-Trunk10] **lacp preempt delay** 20	配置当前Eth-Trunk接口的LACP抢占延时，取值范围为10～180的整数秒 缺省情况下，LACP抢占等待时间为30s，可用**undo lacp preempt delay**命令恢复抢占等待时间为缺省值

（续表）

配置任务	步骤	命令	说明
（可选）配置接收LACP报文超时时间	17	**lacp timeout** { **fast** | **slow** } 例如：[HUAWEI-Eth-Trunk10]**lacp timeout fast**	配置LACP模式下成员接口接收LACP协议报文的超时时间，如果在指定周期内没有收到对端发回的LACP协议确认报文，则会重发原来的LACP协议报文。命令中的选项说明如下。 • **fast**：二选一选项，指定接收报文的超时时间为3s • **slow**：二选一选项，指定接收报文的超时时间为90s 配置此命令后，本端将接收报文的超时时间通过LACP报文通知对端。如本端配置为**fast**接收报文超时时间，则对端在接收到报文后会更改为**fast**的发送周期1s。**两端配置的超时时间可以不一致**，但为了便于维护，建议用户配置一致的LACP协议报文超时时间 缺省情况下，接收报文的超时时间为90s，可用**undo lacp timeout**命令恢复LACP模式下接口接收LACP协议报文的超时时间为缺省值

可以使用 display eth-trunk [trunk-id [interface interface-type interface-number |verbose]] 命令查看Eth-Trunk接口的配置信息；使用display trunkmembership eth-trunk trunk-id命令查看指定编号Eth-Trunk接口的成

员接口信息。

4.5.7 LACP模式的链路聚合配置示例

本示例拓扑结构如图4-8所示，在两台Switch设备上配置LACP模式链路聚合组，而且要求两条活动链路具有负载分担的能力，两设备间的链路具有1条冗余备份链路，当活动链路出现故障时，备份链路替代故障链路，保持数据传输的可靠性。

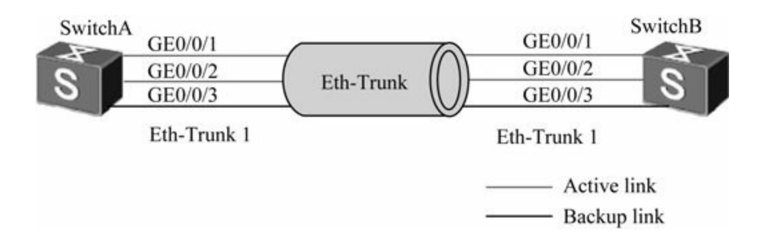


图4-8 LACP模式链路聚合配置示例拓扑结构

1．配置思路分析

因为本示例要求具有链路备份功能，所以只能采用LACP模式的链路聚合方式。根据4.5.6节介绍的九项配置任务，再结合本示例的具体要求可以得出如下基本配置思路。

（1）创建Eth-Trunk，配置Eth-Trunk为LACP模式，实现链路聚合功能。

（2）将3个成员接口GE0/0/1～0/0/3加入Eth-Trunk接口中。

（3）根据两台设备的主次程度配置两台设备的系统LACP优先级，确定主动端，这样就会按照主动端设备的接口选择活动接口。

（4）配置活动接口上限阈值（本示例为2），实现保证带宽的情况下提高网络的可靠性。

（5）配置两端设备中成员接口的LACP优先级，确定活动链路接口，优先级高的接 口将被选作活动接口。

2．具体配置步骤

同样因为本示例中的配置是对称的，所以在此仅以SwitchA上的配置为例进行介绍。

（1）在SwitchA上创建Eth-Trunk1并配置为LACP模式。

```
<HUAWEI>system-view
[HUAWEI] sysname SwitchA
[SwitchA] interface eth-trunk1
[SwitchA-Eth-Trunk1] mode lacp
[SwitchA-Eth-Trunk1] quit
```

（2）配置SwitchA上的GE0/0/1～0/0/3三个成员接口加入Eth-Trunk1接口中。

```
[SwitchA] interfacegigabitethernet 0/0/1
[SwitchA-GigabitEthernet0/0/1] eth-trunk 1
[SwitchA-GigabitEthernet0/0/1] quit
[SwitchA] interface gigabitethernet 0/0/2
[SwitchA-GigabitEthernet0/0/2] eth-trunk 1
[SwitchA-GigabitEthernet0/0/2] quit
[SwitchA] interface gigabitethernet 0/0/3
```

[SwitchA-GigabitEthernet0/0/3] eth-trunk 1

[SwitchA-GigabitEthernet0/0/3] quit

（3）在SwitchA上配置系统优先级为100，使其成为LACP主动端。此时在SwitchB上的优先级值要大于100（值越大，优先级越低），才确保 SwitchA 成为主动端。SwitchB 端可不用配置，因为系统 LACP 优先级为 32768，优先级远小于 SwitchA上配置的100。

[SwitchA] lacp priority 100

（4）在SwitchA上配置活动接口上限阈值为2

[SwitchA] interface eth-trunk 1

[SwitchA-Eth-Trunk1] max active-linknumber2

[SwitchA-Eth-Trunk1] quit

（5）在SwitchA上配置接口优先级确定活动链路。此时在对端SwitchB上这两条链路的对应端口也要配置高优先级，以确保这两条链路为活动链路。

[SwitchA] interfacegigabitethernet 0/0/1

[SwitchA-GigabitEthernet0/0/1] lacp priority 100

[SwitchA-GigabitEthernet0/0/1] quit

[SwitchA] interface gigabitethernet 0/0/2

[SwitchA-GigabitEthernet0/0/2] lacp priority 100

[SwitchA-GigabitEthernet0/0/2] quit

其他的可选配置在此可不用配置。配置完成后可通过display eth-trunk 1命令查看各设备的 Eth-Trunk 信息，查看链路是否协商成功。如下所示。要注意的是，在 LACP模式中在一端设备上执行本命令后可同时查看本端和对端的成员接口配置信息。

```
[SwitchA] display eth-trunk1
Eth-Trunk1's state information is:
Local:
LAG ID:1    WorkingMode:LACP
Preempt Delay: Disabled    Hash arithmetic: According to SIP-XOR-DIP
System Priority: 100    System ID: 00e0-fca8-0417
Least Active-linknumber: 1    Max Active-linknumber: 2
Operate status: up     Number Of Up Port In Trunk: 2
--------------------------------------------------------------------------------
ActorPortName    Status    PortType PortPri    PortNo PortKey    PortState    Weight
GigabitEthernet0/0/1    Selected    1GE    100    6145    2865    11111100    1
GigabitEthernet0/0/2    Selected    1GE    100    6146    2865    11111100    1
GigabitEthernet0/0/3    Unselect    1GE    32768    6147    2865    11100000    1
Partner:
------------------------------------------------------------------------------
ActorPortName    SysPri    SystemID    PortPri PortNo    PortKey    PortState
GigabitEthernet0/0/1    32768    00e0-fca6-7f85    32768    6145    2609    11111100
GigabitEthernet0/0/2    32768    00e0-fca6-7f85    32768    6146    2609    11111100
GigabitEthernet0/0/3    32768    00e0-fca6-7f85    32768    6147    2609    11110000
```

```
[SwitchB] display eth-trunk 1
Eth-Trunk1's state information is:
Local:
LAG ID:1    WorkingMode:LACP
Preempt Delay: Disabled    Hash arithmetic: According to SIP-XOR-DIP
System Priority: 32768    System ID: 00e0-fca6-7f85
Least Active-linknumber: 1    Max Active-linknumber: 8
Operate status: up    Number Of Up Port In Trunk: 2
--------------------------------------------------------------------------------
ActorPortName    Status    PortType PortPri    PortNo PortKey    PortState    Weight
GigabitEthernet0/0/1    Selected    1GE    32768    6145    2609    11111100 1
GigabitEthernet0/0/2    Selected    1GE    32768    6146    2609    11111100 1
GigabitEthernet0/0/3    Unselect    1GE    32768    6147    2609    11100000 1
Partner:
--------------------------------------------------------------------------------
ActorPortName    SysPri    SystemID    PortPri PortNo    PortKey    PortState
GigabitEthernet0/0/1    100    00e0-fca8-0417    100    6145    2865    11111100
GigabitEthernet0/0/2 100    00e0-fca8-0417    100    6146    2865    11111100
GigabitEthernet0/0/3 100    00e0-fca8-0417    32768    6147    2865    11110000
```

通过以上显示信息可以看到，SwitchA的系统优先级为100，高于SwitchB的系统优先级（为缺省的32768）。Eth-Trunk的成员接口中GigabitEthernet0/0/1、GigabitEthernet0/0/2成为活动接口，处于“Selected”状态，接口GigabitEthernet0/0/3处于“Unselect”状态，同时实现M条链路的负载分担和N条链路的冗余备份功能。

4.6 Eth-Trunk接口本地流量优先转发

在设备堆叠或者CSS（Cluster Switch System，集群交换机系统）情况下（有关交换机堆叠和集群将在下章介绍），为了保证流量的可靠传输，流量的出接口通常设置为Eth-Trunk接口。那么Eth-Trunk接口中可能存在跨成员交换机的成员接口。当堆叠设备转发流量时，Eth-Trunk 接口通过 HASH 算法可能会选择跨成员交换机的成员接口，由此增加了跨设备之间的带宽承载压力，也降低了流量转发效率。

为了解决上述问题，可通过本命令使能 Eth-Trunk 接口本地流量优先转发功能，即从本地进入的流量优先通过本地交换机的成员接口转发。如果本地交换机上没有Eth-Trunk的成员接口，再从跨成员交换机的成员接口转发。这样可以有效地节省设备间通信带宽，提高流量转发效率。

4.6.1 使能Eth-Trunk接口本地流量优先转发功能

要使在CSS设备中从本地交换机进入的流量优先通过本地的成员接口转发，就必须使能跨设备 Eth-Trunk 接口上的本地流量优先转发功能。这样也可以减少集群设备之间的带宽承载压力，提高流量转发效率。

使能Eth-Trunk接口本地流量优先转发功能后，当Eth-Trunk接口本地交换机上有出接口且出无故障时，本地的 Eth-Trunk 转发表中将只包含本地交换机的出接口。这样在通过HASH算法选择出接口时只能选中本

地交换机上的接口，流量从本地交换机转发出去。而当Eth-Trunk接口本地交换机上无出接口或者全部故障时，本地交换机的Eth-Trunk转发表中将包含Eth-Trunk接口中所有可转发的出接口。这样在通过HASH算法选择出接口时将选中其他成员交换机上的出接口，流量将通过跨设备转发。

当然，实际情况下不是必须启用这项功能，要根据实际情况配置使能或去使能该功能。

（1）如果本设备 Eth-Trunk 的活动接口的带宽足以承载本设备转发的流量，可以使能 Eth-Trunk 接口本地流量优先转发功能，避免转发效率低、集群设备之间的带宽承载压力大的问题。

（2）如果本设备 Eth-Trunk 的活动接口的带宽不能承载本设备转发的流量，需要去使能 Eth-Trunk 接口本地流量优先转发功能，此时本设备的部分流量就会选择跨设备的Eth-Trunk出接口转发，防止发生丢包。

在配置Eth-Trunk接口本地流量优先转发功能之前，需要确保Eth-Trunk接口已经创建，并已经加入物理接口，当然还必须已经搭建好设备集群环境，同时要确保本设备Eth-Trunk出接口的带宽足以承载本设备转发的流量，防止发生丢包。

配置Eth-Trunk接口本地流量优先转发功能的方法很简单，只需在对应的Eth-Trunk接口视图下执行local-preference enable命令即可。缺省情况下，已经使能了Eth-Trunk接口本地流量优先转发功能，可用 undo local-preference enable命令去使能 Eth-Trunk接口本地流量优先转发功能。

但要注意的是，本地流量优先转发功能只对已知单播有效，不对未知单播生效。

4.6.2 Eth-Trunk接口本地流量优先转发配置示例

本示例拓扑结构如图4-9所示，为了增加设备的容量采用设备堆叠技术，将Switch3和Switch4通过专用的堆叠电缆连接起来，对外呈现为一台逻辑交换机。

为了实现设备间的备份、提高可靠性，采用跨堆叠设备 Eth-Trunk 接口技术，将不同设备上的物理接口加入同一个Eth-Trunk接口。在网络无任何故障情况下，在PE设备上查看成员口信息时，发现VLAN2的数据流量会同时通过GE1/0/1和GE1/0/2成员接口转发（最好是仅从GE1/0/1接口转发），VLAN3的数据流量也会同时通过GE1/0/1和GE1/0/2成员接口转发（最好是仅从GE1/0/2接口转发）。增加了堆叠设备之间的带宽承载能力（因为如果从对端设备接收的本端 VLAN 数据时，需要在堆叠设备之间进行传输），也降低了流量转发效率。

为了有效保证VLAN2的数据流量通过成员口GE1/0/1转发，VLAN3的数据流量通 过成员口GE1/0/2转发，可在堆叠设备上使能Eth-Trunk接口本地流量优先转发功能。

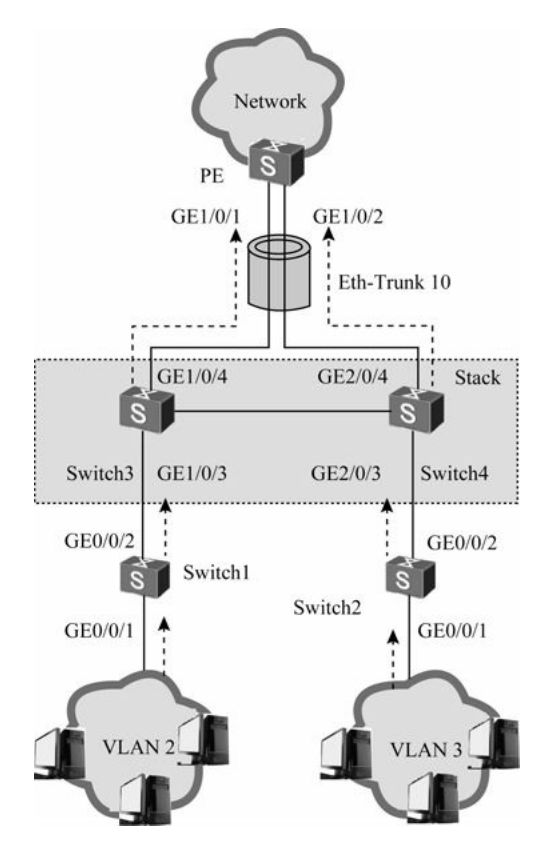


图4-9 Eth-Trunk接口本地流量优先转发配置示例拓扑结构

1．配置思路分析

本示例虽然目的是要启用堆叠设备中Eth-Trunk 接口的本地流量优先转发功能，但因为涉及Eth-Trunk链路聚合，所以必须配置它。同时本示例又涉及S系列交换机的堆叠功能配置，所以本示例中又将介绍交换机堆叠的基本配置，但不作具体介绍，有关交换机堆叠的具体配置方法参见下章相关内容。本示例的基本配置思路如下。

（1）在堆叠交换机和PE交换机上分别创建Eth-Trunk接口（采用缺省的手工负载分担方式），并加入Eth-Trunk的成员接口。

（2）配置交换机堆叠和Switch1和Switch2上的其他接口加入相应VLAN，实现二层互通。有关VLAN的具体配置方法参见本书第6章。

（3）使能Eth-Trunk接口本地流量优先转发功能。

2．具体配置步骤

（1）在交换机堆叠和PE交换机上分别创 建Eth-Trunk接口，并配置为Trunk端口类型，允许通过所有VLAN。

交换机堆叠上的**Eth-Trunk**接口配置：

```
<HUAWEI> system-view
[HUAWEI] sysname Stack
[Stack] interface eth-trunk10
[Stack-Eth-Trunk10] port link-type trunk    #---设置Eth-Trunk接口为Trunk类型
```

[Stack-Eth-Trunk10] port trunk allow-pass vlan all   #---设置Eth-Trunk接口允许所有VLAN报文通过
[Stack-Eth-Trunk10] quit

PE上的**Eth-Trunk**接口配置：

<HUAWEI> system-view
[HUAWEI] sysnamePE
[PE] interface eth-trunk 10
[PE-Eth-Trunk10] port link-type trunk
[PE-Eth-Trunk10] port trunk allow-pass vlan all
[PE-Eth-Trunk10] quit

（2）把交换机堆叠和PE交换机上的对应成员接口加入到它们的Eth-Trunk接口中。

交换机堆叠上的配置：

[Stack] interfacegigabitethernet 1/0/4
[Stack-GigabitEthernet1/0/4] eth-trunk 10
[Stack-GigabitEthernet1/0/4] quit
[Stack] interface gigabitethernet 2/0/4
[Stack-GigabitEthernet2/0/4] eth-trunk10
[Stack-GigabitEthernet2/0/4] quit

PE交换机上的配置：

[PE] interfacegigabitethernet 1/0/1
[PE-GigabitEthernet1/0/1] eth-trunk 10
[PE-GigabitEthernet1/0/1] quit
[PE] interface gigabitethernet 1/0/2
[PE-GigabitEthernet1/0/2] eth-trunk10
[PE-GigabitEthernet1/0/2] quit

（3）配置交换机堆叠、Switch1 和 Switch2 上各接口的 Trunk 类型及所允许通过的VLAN。

交换机堆叠上的配置：

[Stack] vlan batch 2 3
[Stack] interfacegigabitethernet 1/0/3
[Stack-GigabitEthernet1/0/3] port link-type trunk
[Stack-GigabitEthernet1/0/3] port trunk allow-pass vlan 2
[Stack-GigabitEthernet1/0/3] quit
[Stack] interface gigabitethernet 2/0/3
[Stack-GigabitEthernet2/0/3] port link-type trunk
[Stack-GigabitEthernet2/0/3] port trunk allow-pass vlan 3
[Stack-GigabitEthernet2/0/3] quit

Switch1上的配置：

<HUAWEI> system-view
[HUAWEI] sysname Switch1
[Switch1] vlan 2
[Switch1-vlan2] quit

[Switch1] interfacegigabitethernet0/0/1

[Switch1-GigabitEthernet0/0/1] port link-type trunk

[Switch1-GigabitEthernet0/0/1] port trunk allow-pass vlan 2

[Switch1-GigabitEthernet0/0/1] quit

[Switch1] interface gigabitethernet 0/0/2

[Switch1-GigabitEthernet0/0/2] port link-type trunk

[Switch1-GigabitEthernet0/0/2] port trunk allow-pass vlan 2

[Switch1-GigabitEthernet0/0/2] quit

Switch2上的配置：

<HUAWEI> system-view

[HUAWEI] sysname Switch2

[Switch2] vlan 3

[Switch2-vlan3] quit

[Switch2] interface gigabitethernet0/0/1

[Switch2-GigabitEthernet0/0/1] port link-type trunk

[Switch2-GigabitEthernet0/0/1] port trunk allow-pass vlan 3

[Switch2-GigabitEthernet0/0/1] quit

[Switch2] interfacegigabitethernet 0/0/2

[Switch2-GigabitEthernet0/0/2] port link-type trunk

[Switch2-GigabitEthernet0/0/2] port trunk allow-pass vlan 3

[Switch2-GigabitEthernet0/0/2] quit

（4）在交换机堆叠上使能Eth-Trunk10接口的本地流量优先转发功能。

[Stack] interface eth-trunk10

[Stack-Eth-Trunk10] local-preference enable

[Stack-Eth-Trunk10] quit

因为缺省情况下，本地流量优先转发功能处于使能状态，如果以前没有关闭该项功能，此时执行 local-preference enable命令将会提示“Error: The local preferential forwarding mode has been configured.”。不管这个提示信息。

上述配置成功后，在交换机堆叠和PE交换机任意视图下执行display trunkmembership eth-trunk命令可以看到Eth-Trunk接口的成员口信息。如下所示的是在交换机堆叠上执行本命令后的输出信息。

<Stack>display trunkmembership eth-trunk10

Trunk ID: 10

Used status: VALID

TYPE: ethernet

Working Mode : Normal

Number Of Ports in Trunk = 2

Number Of Up Ports in Trunk = 2

Operate status: up

Interface GigabitEthernet1/0/4, valid, operate up, weight=1

Interface GigabitEthernet2/0/4, valid, operate up, weight=1

4.7 E-Trunk

E-Trunk是一种实现跨设备链路聚合的控制协议，基于LACP（单台设备链路聚合的标准）进行了扩展，能够实现多台设备间的链路聚合。从而把链路可靠性从单板级提高到了设备级。本功能除**S2700**、**S5700LI**和**S5700S-LI**系列外，其他支持VRP系统的华为S系列交换机均支持。

有关E-Trunk方面基础知识介绍参见本章4.5.1节。

在配置E-Trunk之前，需要完成以下任务：

（1）正确连接设备之间的物理链路。

（2）配置静态LACP模式Eth-Trunk接口。

4.7.1 E-Trunk配置任务

E-Trunk的配置任务如下。

1．配置E-Trunk的LACP系统ID和优先级

在E-Trunk中，为了使CE设备认为对端的两台PE设备是一台设备，E-Trunk中主、备两台PE设备的LACP优先级、系统ID都需要保持一致。

2．创建E-Trunk并配置优先级

E-Trunk的优先级用于在聚合组中决策两台设备的主备状态。

3．配置本端和对端的IP地址

E-Trunk协议报文采用本端配置的源IP地址及协议端口号发送。但如果要修改地址则两台设备需要同时修改，否则会导致协议报文丢弃。

4．配置E-Trunk与BFD会话绑定

通过报文接收超时无法快速感知对端是否故障，可以使用快速检测协议BFD快速感知。每个E-Trunk都需要指定对端的IP，通过创建检测对端路由是否可达的BFD会话， E-Trunk可感知到BFD通告的故障，并快速处理。

5．将Eth-Trunk加入E-Trunk

当E-Trunk配置成功，必须向E-Trunk中加入成员Eth-Trunk，才能实现两台设备上的链路聚合协议。从而实现跨设备的链路聚合组冗余，提高网络可靠性。

6．（可选）配置Eth-Trunk在E-Trunk中的工作模式

只能对已经加入E-Trunk的Eth-Trunk接口配置工作模式，Eth-Trunk的工作模式分为自动模式、强制主用模式和强制备用模式。强制主用模式就是强制对应 Eth-Trunk 接口为主用状态；强制备用模式就是强制对应 Eth-Trunk 接口为备用状态；自动模式就是根据协商，自动选择工作模式。

当设置工作模式为自动模式或者工作模式由强制模式切换为自动模式后，根据本端E-Trunk的主备状态和对端Eth-Trunk的故障信息决定本端成员Eth-Trunk的状态。

若本端 E-Trunk 状态为主用，则本端 Eth-Trunk 的工作模式为主用。若本端E-Trunk 状态为备用，则对端成员 Eth-Trunk 为故障，则本端 Eth-Trunk 的工作模式为主用。当本端收到对端Eth-Trunk故障恢复消息后，该对端Eth-Trunk进入备用状态。

7．（选）配置密码

为了提高系统的安全性可配置加密密码，对通过E-Trunk的报文进行加密，以确保在E-Trunk通信中Eth-Trunk接口只有收到密码一致的报文才可接收。E-Trunk中的两端设备上的加密密码必须配置为一致。

用户可以选择采用明文加密或密文加密。使用明文加密时，在配置文件中采用明文形式显示；使用密

文加密时，在配置文件中采用加密后的乱码显示，不显示真正的密码，更加安全。

8．（可选）配置超时时间

如果处于备用状态的 E-Trunk 在超时时间内没有收到对端发送的 Hello 报文，则在定时器超时后进入主用状态。此处的超时时间是对端报文中所携带的超时时间，而不是本端设置的超时时间。

9．（可选）配置延时回切时间

当E-Trunk的本端设备处于主用状态时，由于其中某个成员Eth-Trunk的物理状态变为Down，经过LACP协商后对端的成员Eth-Trunk的物理状态变为Up。此时，对端设备变为主用状态，本端设备变为备用状态。当本端故障消除，经过LACP协商，本端恢复为主用状态。

当E-Trunk与其他业务配合使用时，如果E-Trunk状态为主用的设备发生故障恢复后，成员Eth-Trunk状态恢复早于其他相关业务恢复。如果马上将E-Trunk成员的流量回切，会导致业务流量中断。配置 E-Trunk的延时回切时间后，必须等待延时回切定时器超时，本端成员Eth-Trunk状态才能Up，E-Trunk的本端设备才能恢复为主用状态。从而延迟了 E-Trunk成员的流量回切时间，保证业务流量不会中断。

4.7.2 E-Trunk配置与管理

上节介绍的E-Trunk九项主要配置任务的具体配置步骤如表4-19所示。

表4-19 E-Trunk的配置与管理步骤

配置任务	步骤	命令	说明
配置 E-Trunk 的 LACP 系统 ID 和优先级	1	**system-view** 例如：<HUAWEI> **system-view**	进入系统视图
	2	**lacp e-trunk system-id** *mac-address* 例如：[HUAWEI] **lacp e-trunk system-id** 00E0-FC00-0000	配置 E-Trunk 的 LACP 系统 ID，格式为 H-H-H，其中 H 为 4 位的十六进制数，可以输入 1～4 位，如 00e0、fc01。当输入不足 4 位时，表示前面的几位为 0，如输入 e0，等同于 00e0。系统 ID 不能为全 0 或全 F。 【注意】**E-Trunk 中主备两台设备的 LACP 系统 ID 需要保持一致**。当设备上配置多个 E-Trunk 时，不同聚合组的 LACP 系统 ID 可以不同，但此时需要在 Eth-Trunk 接口视图下配置 LACP 系统 ID。在系统视图下配置的 LACP 系统 ID 对所有 Eth-Trunk 接口有效；在 Eth-Trunk 接口视图下配置的 LACP 系统 ID 仅对该 Eth-Trunk 接口有效。对于指定的 Eth-Trunk 接口，如果已经在系统视图下执行了本命令，又在指定的 Eth-Trunk 接口视图下执行本命令，则以 Eth-Trunk 接口视图下配置的值为准 缺省情况下，使用主控板的以太口 MAC 地址作为 E-Trunk 的 LACP 系统 ID。在 Eth-Trunk 接口视图下的缺省值为系统视图下的当前值，可用 **undo lacp e-trunk system-id** 命令删除 E-Trunk 的 LACP 系统 ID
	3	**lacp e-trunk priority** *priority* 例如：[HUAWEI] **lacp e-trunk priority** 10	配置 E-Trunk 的 LACP 优先级，取值范围为 0～65 535 的整数。**值越小 LACP 优先级越高**。如果配置了 LACP 优先级，E-Trunk 中的成员 Eth-Trunk 端口发送 LACP 报文时，采用配置的优先级。否则，使用 E-Trunk 的 LACP 优先级缺省值为 32768 【注意】**E-Trunk 中主备两台设备的 LACP 优先级需要保持一致**。当设备上配置多个 E-Trunk 时，不同聚合组的 LACP 优先级可以不同，此时需要在 Eth-Trunk 接口视图下配置 LACP 优先级。在系统视图下配置的 LACP 优先级对所有 Eth-Trunk 接口有效；在 Eth-Trunk 接口视图下配置的 LACP 优先级仅对该 Eth-Trunk 接口有效。对于指定的 Eth-Trunk 接口，如果已经在系统视图下执行了本命令，又在指定的 Eth-Trunk 接口视图下执行本命令，则以 Eth-Trunk 接口视图下配置的值为准 缺省情况下，E-Trunk 的 LACP 优先级是 32768，可用 **undo lacp e-trunk priority** 命令删除 E-Trunk 的 LACP 优先级
创建 E-Trunk 并配置优先级	4	**e-trunk** *e-trunk-id* 例如：[HUAWEI] **e-trunk** 2	创建 E-Trunk，指定 E-Trunk 编号，取值范围为 1～16 的整数。当 E-Trunk 存在时，执行本命令直接进入 E-Trunk 视图 **在一个 E-Trunk 内，两端设备上配置的 E-Trunk 编号必须相同**。缺省情况下，没有创建任何 E-Trunk，可用 **undo e-trunk** *e-trunk-id* 命令删除指定的 E-Trunk

（续表）

配置任务	步骤	命令	说明
创建 E-Trunk 并配置优先级	5	**priority** *priority* 例如：[HUAWEI-e-trunk-2]**priority** 10	配置 E-Trunk 的优先级，取值范围为 1～254 的整数。优先级用于两台设备间进行主备协商，优先级高的为主用设备，值越小优先级越高。如果优先级相同，那么比较两台设备的系统 ID，ID 较小的为主用设备。如果优先级和系统 ID 都相同，则认为配置错误，丢弃报文，不作处理 缺省情况下，E-Trunk 的优先级为 100，可用 **undo priority** 命令恢复 E-Trunk 的优先级为缺省的 100
配置本端和对端的 IP 地址	6	**peer-address** *peer-ip-address* **source-address** *source-ip-address* 例如：[HUAWEI-e-trunk-2] **peer-address 2.2.2.2 source-address 1.1.1.1**	配置对端和本端的 IP 地址。命令中的参数说明如下： （1）*peer-ip-address*：指定对端 IP 地址 （2）*source-ip-address*：指定对端源 IP 地址 可用 **undo peer-address** 命令删除 E-Trunk 的对端和本端的 IP 地址
配置 E-Trunk 与 BFD 会话绑定	7	**e-trunk track bfd-session session-name** *bfd-session-name* 例如：[HUAWEI-e-trunk-2]**e-trunk track bfd-session session-name** e-trunk-bfd	绑定 BFD 会话，参数 *bfd-session-name* 用来指定要绑定的 BFD 会话名称，为 1～15 个字符，支持空格，不区分大小写。BFD 用于实现 E-Trunk 的两台设备之间控制协议链路的快速故障检测 缺省情况下，E-Trunk 没有绑定 BFD 会话，可用 **undo e-trunk track bfd-session** 命令取消绑定的 BFD 会话
将Eth-Trunk 加入 E-Trunk	8	**quit** 例如：[HUAWEI-e-trunk-2] **quit**	退出 E-Trunk 视图，返回系统视图
	9	**interface eth-trunk** *trunk-id* 例如：[HUAWEI] **interface eth-trunk** 1	进入要加入到 E-Trunk 的 Eth-Trunk 接口视图。但仅 LACP 模式的 Eth-Trunk 才能加入 E-Trunk
	10	**e-trunk** *e-trunk-id* [**remote-eth-trunk** *eth-trunk-id*] 例如：[HUAWEI-eth-trunk1] **e-trunk** 2	将以上 Eth-Trunk 加入到指定 E-Trunk 中。参数说明如下。 （1）*e-trunk-id*：指定以上 Eth-Trunk 接口要加入的 E-Trunk 编号，取值范围为 1～16 的整数 （2）*eth-trunk-id*：可选参数，指定远端 PE 设备的 Eth-Trunk 编号，取值范围为 0～4 294 967 295 的整数 【说明】一个 Eth-Trunk 只能加入一个 E-Trunk。一个 E-Trunk 中，两端设备上所加入的 Eth-Trunk ID 可以不一致，当两台 PE 设备上创建的 Eth-Trunk ID 不一样，如果用户通过本命令将两端PE 设备上不同 ID 的 LACP 模式的 Eth-Trunk 加入同一个 E-Trunk 时，必须选择 **remote-eth-trunk** 参数指定远端 Eth-Trunk ID，能保证 E-Trunk 正常工作 可用 **undo e-trunk** 命令删除指定 E-Trunk 中的 Eth-Trunk
（可选）配置 Eth-Trunk 在 E-Trunk 中的工作模式	11	**quit** 例如：[HUAWEI-eth-trunk1] **quit**	退出 Eth-Trunk 接口视图
	12	**e-trunk** *e-trunk-id* 例如：[HUAWEI] **e-trunk** 2	进入前面创建的 E-Trunk 视图

（续表）

配置任务	步骤	命令	说明
（可选）配置Eth-Trunk在E-Trunk中的工作模式	13	**e-trunk mode** { **auto** | **force-master** | **force-backup** } 例如：[HUAWEI-e-trunk-2]**e-trunk mode force-master**	配置 Eth-Trunk 在 E-Trunk 中的工作模式。选项说明如下。 （1）**auto**：多选一选项，指定 Eth-Trunk 的工作模式为自动模式。当设置工作模式为自动模式或者工作模式由强制模式切换为自动模式后，根据本端 E-Trunk 的主备状态和对端 Eth-Trunk 的故障信息决定本端成员 Eth-Trunk 的状态 （2）**force-master**：多选一选项，指定 Eth-Trunk 的工作模式为强制主用状态。若本端 E-Trunk 状态为主用，则本端 Eth-Trunk 的工作模式为主用；若本端 E-Trunk 状态为备用，对端成员 Eth-Trunk 为故障，则本端 Eth-Trunk 的工作模式为主用 （3）**force-backup**：多选一选项，指定 Eth-Trunk 的工作模式为强制备用状态。当本端收到对端 Eth-Trunk 故障恢复消息后，该 Eth-Trunk 进入备用状态 **只能对已经加入 E-Trunk 的 Eth-Trunk 执行本命令。当 Eth-Trunk 退出 E-Trunk 时，该配置将自动清除** 缺省情况下，Eth-Trunk 在 E-Trunk 中工作在自动模式，可用 **undo e-trunk mode** 命令恢复 Eth-Trunk 在 E-Trunk 中的工作模式为缺省的自动模式
（可选）配置密码	14	**security-key** { **simple** *simple-key* | **cipher** *cipher-key* } 例如：[HUAWEI-e-trunk-2]**security-key cipher** 00E0FC000000	配置加密报文的密码。命令中的参数说明如下。 （1）*simple-key*：二选一参数，指定以明文方式加密安全密钥，为 1～255 整数个字符，不支持空格、单引号和问号，区分大小写。缺省值是 00E0FC0000000000 （2）*cipher-key*：二选一参数，指定以密文方式加密安全密钥，字符串形式，不支持空格、单引号和问号，区分大小写。此时输入密码有两种方式，一种是明文，一种是密文。当输入明文密码时，长度范围为 1～255 整数个字符；当输入密文密码时，长度为 32～392 整数个字符 缺省情况下，simple 方式密码为 00E0FC0000000000，可用 **undo security-key** 命令恢复密码为缺省值
（可选）配置超时时间	15	**timer hello** *hello-times* 例如：[HUAWEI-e-trunk-2] **timer hello** 9	设置主备交换机发送的 Hello 报文时间间隔，备用交换机经过下一步 **timer hold-on-failure multiplier** *multiplier* 命令中参数 *multiplier* 值个发送周期没收到 Hello 报文则会进入主用状态，取值范围为 5～100 的整数，单位是 100ms 缺省情况下，Hello 报文发送周期值为 10，单位为 100ms，即 1s，可用 **undo timer hello** 命令恢复 Hello 报文发送周期为缺省值
	16	**timer hold-on-failure multiplier** *multiplier* 例如：[HUAWEI-e-trunk-2] **timer hold-on-failure multiplier** 2	配置检测 Hello 报文的时间倍数，取值范围为 3～300 的整数。超时时间=发送周期×时间倍数。建议将时间倍数设置为 3 倍以上 对端利用接收到的报文中携带的超时时间来检测本端是否超时。如果对端处于备用状态，在超时时间内没有收到由本端发送的 Hello 报文，则在定时器超时后对端设备进入主用状态 缺省情况下，检测 Hello 报文的时间倍数为 20，可用 **undo timer hold-on-failure multiplier** 命令恢复为缺省值

（续表）

配置任务	步骤	命令	说明
（可选）配置延时回切时间	17	**timer revert delay** *delay-value* 例如：[HUAWEI-e-trunk-2] **timer revert delay** 20	配置回切延迟时间，取值范围为 0～3 600 的整数秒 当 E-Trunk 与其他业务配合使用时，如果 E-Trunk 状态为主用的设备发生故障恢复后，成员 Eth-Trunk 状态恢复早于其他相关业务恢复。执行本命令配置 E-Trunk 的延时回切时间后，必须等待延时回切定时器超时，本端成员 Eth-Trunk 状态才能 Up，E-Trunk 的本端设备才能恢复为主用状态。从而延迟了 E-Trunk 成员的流量回切时间，保证业务流量不会中断 缺省情况下，E-Trunk 延时回切的时间为 120s，可用 **undo timer revert delay** 命令恢复 E-Trunk 延时回切的时间为缺省值

可用display e-trunk e-trunk-id命令查看指定编号的E-Trunk的配置信息。

4.8 Eth-Trunk子接口配置与管理

当二层网络中的交换设备划分到不同的 VLAN 中，为了保证不同 VLAN 间的用户正常通信，需要在三层设备与二层设备相连的 Eth-Trunk 接口上创建子接口与下游用户的VLAN分别对应，并在子接口上配置IP地址。

如图4-10所示，PE1和PE2均为三层交换设备，它们分别与对端的CE1和CE2建立了Eth-Trunk连接，在PE1和PE2端的Eth-Trunk接口上划分Eth-Trunk子接口。目前只有S5710EI、S5700HI、S7700、S9300和S9700系列交换机支持Eth-Trunk子接口。

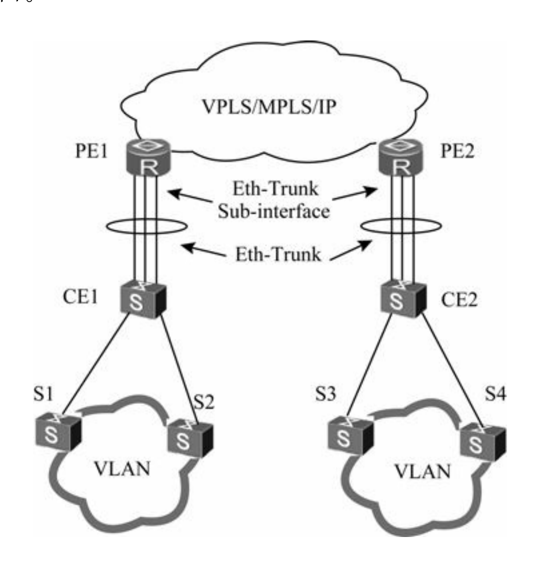


图4-10 Eth-Trunk子接口应用示例

在Eth-Trunk子接口上封装802.1Q并关联VLAN后，VLAN可以通过Eth-Trunk子接口与 VLAN 外的设备通信。Eth-Trunk 子接口与本章前面介绍的以太网子接口一样，也应用于Dot1q终结、QinQ终结等场合。使用二层Eth-Trunk子接口后，Eth-Trunk主接 口上运行二层功能，Eth-Trunk子接口运行三层功能。

Eth-Trunk子接口的配置与管理方法如表4-20所示。主要是在Eth-Trunk接口上创建子接口，在子接口上配置IP地址，进行VLAN终结封装和使能ARP广播功能，总体上与4.4.1节表4-10中介绍的以太网子接口的配置方法基本一样。

表4-20 Eth-Trunk子接口的配置步骤

步骤	命令	说明	
1	**system-view** 例如：< HUAWEI > **system-view**	进入系统视图	
2	**interface eth-trunk** *trunk-id.subnumber* 例如：[HUAWEI] **interface eth-trunk** 2.1	在对应 Eth-Trunk 接口上创建指定的子接口，进入 Eth-Trunk 子接口视图。参数用来指定所创建的 Eth-Trunk 子接口编号，取值范围为 1～4 096	
3	**ip address** *ip-address* { *mask* | *mask-length* } [**sub**] 例如：[HUAWEI-Eth-Trunk2.1]**ip address** 192.168.0.10 255.255.255.0	为以上 Eth-Trunk 子接口配置 IP 地址。缺省情况下，在子接口上没有配置 IP 地址，可用 **undo ip address** *ip-address* { *mask* | *mask-length* } [**sub**] 删除子接口上指定的 IP 地址	
4	**dot1q termination vid** *low-pe-vid* [**to** *high-pe-vid*] 例如：[HUAWEI-Eth-Trunk2.1]**dot1q termination vid** 4	（二选一）单层或双层 VLAN 标签终结	配置 Eth-Trunk 子接口对一层 Tag 报文的终结功能。缺省情况，子接口没有配置 dot1q 封装的单层 VLAN ID，可用 **undo dot1q termination vid** *low-pe-vid* [**to** *high-pe-vid*] 命令取消子接口 dot1q 封装的单层 VLAN ID 其他说明参见 4.4.1 节表 4-10 中的第 4 步
	qinq termination pe-vid *pe-vid* **ce-vid** *ce-vid1* [**to** *ce-vid2*] 例如：[[HUAWEI-Eth-Trunk2.1]**qinq termination pe-vid** 4 **ce-vid** 10 **to** 12		配置 Eth-Trunk 子接口对两层 Tag 报文的终结功能。缺省情况，子接口没有配置 QinQ 封装的双层 VLAN ID，可用 **undo qinq termination pe-vid ce-vid** 命令取消子接口 QinQ 封装的双层 VLAN ID 其他说明参见 4.4.1 节表 4-10 中的第 4 步
5	**arp broadcast enable** 例如：[HUAWEI-Eth-Trunk2.1] **arp broadcast enable**	使能 Eth-Trunk 子接口的 ARP 广播功能。缺省情况下，终结子接口没有使能 ARP 广播功能 **undo arp broadcast enable** 命令去使能终结子接口的 ARP 广播功能 其他说明参见 4.4.1 节表 4-10 中的第 5 步	

可用以下命令查看Eth-Trunk子接口上的相关配置：

（1）display interface eth-trunk [trunk-id [.subnumber]]：查看指定Eth-Trunk子接口的状态。

（2）display dot1q information termination [interface eth-trunk [trunk-id [.subnumber]]]：查看配置了dotlq终结的所有接口的名称以及终结子接口对用户报文终结的规则数量。

（3）display qinq information termination [interface i eth-trunk [trunk-id [.subnumber]]]：查看配置了QinQ终结的所有接口的名称以及终结子接口对用户报文终结的规则数量。

第5章　交换机堆叠和集群配置与管理

交换机堆叠和集群是两种可解决单台交换机性能不足、容易出现单点故障问题的交换机管理技术。实现堆叠和集群后的各成员交换机一起可看成一台逻辑交换机系统，可通过一个IP地址进行管理，不仅大大简化对各成员交换机的管理，而且从整体上提高单台交换机的性能，其中的各成员交换机间还可实现负载均衡和容错，避免因单台交换机出现故障而出现网络中断。

在华为S交换机中也支持交换机堆叠和集群功能，其中堆叠技术为iStack（Intelligent Stack，智能堆叠），S2700、S3700、S5700和S6700中低端系列交换机支持；集群技术称之为CSS（Cluster Switch System，集群交换系统）， S7700、S9300、S9300E和S9700高端交换机系列支持。但是不同交换机系列所支持的堆叠或者集群连接方式有所不同（有通过专门的堆叠卡或集群卡连接的，有可通过普通业务口连接的），当然在配置上也有所同不同，在配置时一定要注意。这两种交换机管理技术在工作原理、配置方法和功能支持上存在许多相似性，但华为iStack堆叠中可以支持的成员交换机更多些（最多为9台），目前华为CSS只支持两台交换机的集群。

本章要介绍华为S系列交换机的iStack堆叠和CSS集群配置与管理方法，它们的配置和管理方法都是比较容易的。

5.1 iStack基础

交换机堆叠是一种提高端口可用背板带宽，扩展交换机端口，提高交换机可靠性，集中管理多台交换机十分有效的技术。它可将多台支持堆叠特性的交换机组合在一起，从逻辑上组合成一台整体交换机，不仅可以通过一个命令行界面，一个IP地址进行集中管理，还可以提高单台交换机的转发性能和可靠性，实现各成员交换机间的负载均衡。在如图5-1所示的拓扑结构中，左图最上面的两台交换机通过堆叠功能后就可看成右图最上面那一台交换机，这就是交换机堆叠的最直接外在表现形式。

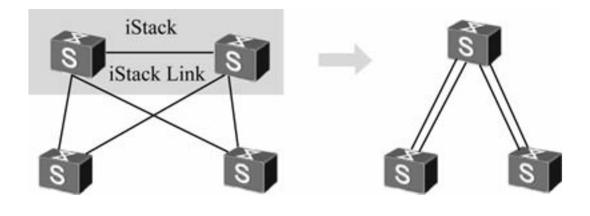


图5-1 iStack堆叠示意图

在华为S系列交换机中，交换机堆叠技术称之为 iStack（Intelligent Stack，智能堆 叠）。它主要适用于需要更高端口密度和交换性能的中低端交换机中，如S2700、S3700、S5700、S6700系列交换机。而高端的S7700、S9300、S9300E和S9700系列则采用的是更高级的性能扩展技术——CSS （集群交换系统），不再支持iStack，具体也将在本章后面介绍。

5.1.1 iStack概述

iStack是华为S系列交换机的堆叠技术。与其他堆叠技术一样，在iStack堆叠系统建立之前，每台成员交换机都是单独的实体，都有自己独立的IP地址和MAC地址，对外体现为多台交换机，用户需要独立的管理所有的交换机；而在iStack堆叠系统建立后，堆叠中的所有成员交换机对外体现为一个统一的逻辑实体，用户使用一个IP地址就可以对堆叠中的所有交换机进行管理和维护。通过交换机堆叠，可以实现网络大数据量转发和网络高可靠性，同时简化网络管理。

1．交换机角色

在iStack堆叠中所有的单台交换机都称为成员交换机（最多可以有8台，或者9台，不同系列所支持的成员交换机数不一样，本章后面将具体介绍），但按照各自功能的不同又可以分为以下三种角色。

（1）主交换机。主交换机在交换机堆叠配置文件中显示为Master，负责整个堆叠系统的管理。一个交换机堆叠只有一台主交换机 。

（2）备交换机。备交换机在交换机堆叠配置文件中显示为Standby，是主交换机的备用交换机，用于当原主交换机出现故障时接替原主交换机的工作，管理整个堆叠系统。与主交换机一样，一个交换机堆叠也只有一台备交换机 。

（3）从交换机。从交换机在交换机堆叠配置文件中显示为Slave。在一个交换机堆叠系统中，除了主交换机外的其他所有交换机（包括备交换机 ）都是从交换机。

2．堆叠ID

为了便于识别和管理交换机堆叠中各成员交换机，为所有成员交换机（包括主交换机和备交换机）都分配了一个堆叠 ID，即成员编号（Member ID），且每个成员交换机的 堆叠ID都是唯一的。

3．堆叠优先级

既然前面说了在一个交换机堆叠中有一个主交换机，还有一个备交换机，那么如何确定哪台成员交换机担当这两种比较特殊的角色呢？这就要用到堆叠优先级属性了，用于堆叠角色选举过程中确定主交换机和备交换机角色。优先级值越大表示优先级越高，优先级越高当选为主交换机和备交换机的可能性越大。

4．堆叠物理成员端口

这里所说的“堆叠物理成员端口”是指采用普通业务口堆叠连接方式时，各成员交换机上用于堆叠连接的物理业务端口，不是指堆叠卡上的专门堆叠接口。堆叠物理成员端口用于转发需要跨成员交换机的业务报文或成员交换机之间的堆叠协议报文。

5．堆叠端口

这里所说的“堆叠端口”是指采用普通业务口堆叠连接方式时用于堆叠连接的逻辑端口，需要和上面介绍的堆叠物理成员端口绑定，即向逻辑堆叠端口中添加物理成员端口。堆叠的每台成员交换机上支持两个堆叠端口：Stack-Portn/1和Stack-Portn/2，其中n为成员交换机的堆叠ID。

5.1.2 iStack特性的产品支持

在最新的Sx700大系中，只有S2700、S3700、S5700和S6700系列支持iStack堆叠功能，但这些系列中也并不是所有机型都支持，而且S2700/3700系列与S5700/6700系列对iStack的特性支持也不完全一样。

在 S2700/3700 系列中支持堆叠的交换机子系列有 S3700EI、S3700SI、S2700–52P-EI、S2700–52P-PWR-EI和S2710SI，但不同子系列的交换机不能混合堆叠 。其中S3752EI和S3752SI子系列最多支持8台交换机堆叠，其他子系列产品最多支持9台交换机堆叠。需要特别说明的是：其中的S3752EI子系列交换机不能与S3728EI子系列交换机组成堆叠；S3752SI子系列交换机也不能与S3728SI子系列交换机组成堆叠。也就

是堆叠中各成员交换机不仅功能版本一样，而且一般来说端口数也一样。

S5700和S6700系列中目前支持9台相同型号交换机组成堆叠，不同型号交换机之间也不可以混合堆叠。不同机型所支持的堆叠连接方式将在本节后面介绍。

1．堆叠主、备交换机选举

iSack 交换机堆叠系统由多台成员交换机组成，每台成员交换机具有一个确定的角色。堆叠建立时，成员交换机间相互发送堆叠竞争报文，选举出主、从交换机。当从交换机的VRP系统软件版本号与主交换机的VRP系统软件版本不一致时，从交换机将自动同步 主交换机的VRP系统软件版本，复位重启后加入堆叠系统。主交换机收集成员信息并计算堆叠拓扑，然后将堆叠拓扑信息同步到所有的成员交换机。

主交换机选举规则如下。

（1）首先进行运行状态比较，已经运行的堆叠交换机中最先处于启动状态的交换机 将被选举为主交换机。

（2）如果有多台成员交换机都已处于启动状态，则再对这些交换机进行堆叠优先级比较，堆叠优先级高的交换机优先选举为主交换机。

（3）如果某些成员交换机的堆叠优先级也一样，则再对这些成员交换机进行 MAC地址比较，MAC地址小的交换机优先选举为主交换机。

备交换机选举规则如下。

（1）除主交换机外其他各成员交换机中最先处于启动状态的交换机成为备份交换机。

（2）如果有多台除主交换机外的其他交换机同时完成启动时，则这些成员交换机中堆叠优先级最高的交换机成为备交换机。

（3）如果以上这些交换机的堆叠优先级也相同，则 MAC 地址最小的将选举为备交换机。

2．堆叠连接方式

不同S系列交换机的iStack堆叠连接方式也不完全一样。**S2700**和**S3700**系列主要支持堆叠卡连接方式，是通过专门的堆叠卡中提供的堆叠端口（也可使用复用上行千兆口作为堆叠端口 ）和专用的SFP高速堆叠电缆（连接器为1.5m长的20针SFP公头，如图5-2所示）连接的（堆叠线缆两端插头需配戴防静电防护帽）。



图5-2 SPF堆叠线缆

S5700和S6700系列支持以下两种堆叠连接方式。

（1）堆叠卡连接：各成员交换机之间通过专用的堆叠卡ETPC和专用的PCI-E堆叠电缆（如图5-3左图所示，长度为1m）连接。仅**S5700EI**和**S5700SI**子系列支持这种连接方式 。

（2）业务口连接：各成员交换机间通过堆叠端口绑定的堆叠物理成员端口和SFP+高速电缆（如图5-3右图所示，连接器为 20 针 SFP+公头，长度可以有1m、3m和10m三种规格）相连，不需要专用的堆叠插卡。仅**S5700LI**、**S5710EI**和**S6700**系列支持这种连接方式 。

S5700和 S6700系列中各子系列所支持的堆叠连接方式和连接性能说明如表 5-1所示。



图5-3 PCI-E堆叠电缆和SFP+堆叠高速电缆

表5-1 S5700和S6700系列中的各子系列所支持的堆叠连接方式

子系列	堆叠方式	支持堆叠的接口	堆叠线缆	最大堆叠带宽（单向）	说明
S5700-P-LI（GE 上行款型）	业务口堆叠	V200R001 版本：交换机最后的 2 个 SFP 接口 V200R002 及以后版本：交换机最后的 4 个 SFP 接口	1m 无源 SFP+电缆 10m 有源 SFP+电缆	使用 1m 无源 SFP+电缆为 2.5Gbit/s 使用 10m 有源 SFP+电缆为 5Gbit/s	V200R001 版本：单交换机最多支持 2 个堆叠口，每个堆叠口最多包含 1 个物理成员口，单交换机最大支持 2 个物理成员口 V200R002 及以后版本：单交换机最多支持 2 个堆叠口，每个堆叠口最多包含 2 个物理成员口，单交换机最大支持 4 个物理成员口 **支持 GE 上行款型之间混堆，不支持 GE 上行款型与 10GE 上行款型之间混堆** **S5700-10P-LI-AC 和 S5700-10P-PWR-LI-AC 不支持堆叠**
S5700-X-LI（10GE 上行款型）	业务口堆叠	交换机最后的 4 个 SFP+接口	1m/3m 无源 SFP+电缆、10m 有源 SFP+电缆、普通的 SFP+光模块和光纤	10Gbit/s	单交换机最多支持 2 个堆叠口，每个堆叠口最多包含 2 个物理成员口，单交换机最大支持 4 个物理成员口 **支持 10GE 上行款型之间混堆，不支持 GE 上行款型与 10GE 上行款型之间混堆**
S5700-SI	堆叠卡堆叠	堆叠卡的两个堆叠口	1m 的 PCIe 电缆	12Gbit/s	**支持 S5700-SI 的所有 PoE 和非 PoE 款型混堆** **S5700-26X-SI-12S-AC 不支持堆叠**
S5700-EI	堆叠卡堆叠	堆叠卡的两个堆叠口	1m/3m 的 PCIe 电缆	12Gbit/s	**只有 S5700-52C-EI 和 S5700-28C-EI-24S 支持，但支持 S5700-EI 的所有 PoE 和非 PoE 款型混堆**
S5710-EI	业务口堆叠	交换机上任意 10GE 接口：包括交换机前面固定的 4 个 SFP+接口和后面的 SFP+插卡	1m/3m 无源 SFP+电缆、10m 有源 SFP+电缆、普通的 SFP+光模块和光纤	10Gbit/s	V200R001 版本：单交换机最多支持 2 个堆叠口，每个堆叠口最多包含 3 个物理成员口，单交换机最大支持 4 个物理成员口。所有堆叠口的物理成员口分布必须全部位于前面板或全部位于后面的插卡上 V200R002 及以后版本：单交换机最多支持 2 个堆叠口，每个堆叠口最多包含 4 个物理成员口，单交换机最大支持 8 个物理成员口 **支持 S5710-EI 的所有款型混堆**
S6700	**业务口堆叠**	交换机上任意 **10GE** 接口。最多同时 **8** 个接口用于堆叠	**1m/3m/10m 无源 SFP+电缆、10m 有源 SFP+电缆（V200R001C00 版本及以后版本支持）、普通 SFP+光模块和光纤**	10Gbit/s	**支持 S6700 的所有款型之间混堆，接口工作在 GE 模式时不支持堆叠**
S5700-HI 和 S5700S-LI 子系列目前暂不支持堆叠					

3．堆叠连接拓扑结构

华为S系列交换机iStack堆叠的连接拓扑结构有“链形连接”和“环形连接”两种。 环形连接拓扑结构是堆叠成员交换机通过堆叠端口交叉相连形成一个“环”形结构，如图5-4所示。

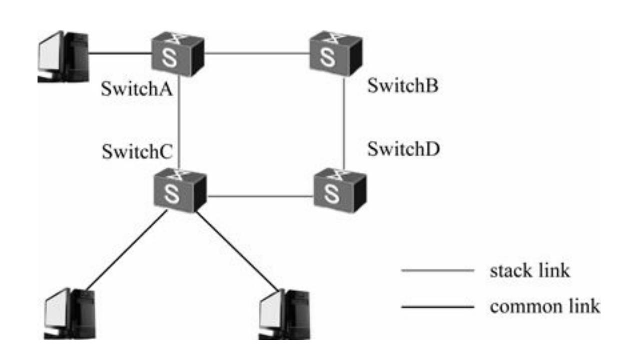


图5-4 iStack堆叠环形拓扑结构

链形连接拓扑结构中处于链两端的交换机只使用一个堆叠端口与邻居交换机相连，最终形成一个“链条”形结构（有点像交换机“级连”），如图5-5所示。相比之下，环形连接拓扑结构比链形连接拓扑结构具有更高的可靠性，因为当 链形连接拓扑结构中出现链路故障时会引起堆叠分裂；而当环形连接拓扑结构中某条链路故障时会形成链形连接，整体堆叠的业务不会受到影响。所以建议在实际部署业务时采用环型拓扑结构部署。

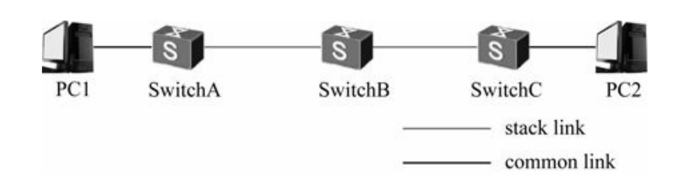


图5-5 iStack堆叠链形拓扑结构

4．堆叠的管理和维护

iStack 堆叠建立后，所有的成员交换机形成一台逻辑交换机存在于网络中，所有成员交换机的资源由堆叠主交换机统一管理。用户可以通过任意一台成员交换机的网管接口或串口登录堆叠系统 ，对整个堆叠系统进行管理和维护。但同一时刻只能由一个网管接口或串口登录 。

另外，在管理堆叠中的成员交换机时与管理单台交换机时在接口编号方面有些区别。对于单台没有运行堆叠的交换机，接口编号采用：0/子卡号/端口号；而在交换机加入堆叠后，接口编号采用：堆叠 ID/子卡号/端口号。如交换机没有运行堆叠时，某个接口的编号为GigabitEthernet0/0/1；在该交换机加入堆叠后，如果为该交换机分配的堆叠ID为2，则该接口的编号将变为GigabitEthernet2/0/1。这样就可以使得整个堆叠交换机中各接口编号具有唯一性。

5．堆叠成员加入

在iStack堆叠维护和使用过程中会继续进行拓扑收集工作，当发现有新的成员交换机（已配置了堆叠连接和堆叠功能）加入时会根据新加入交换机的状态采取不同的处理。

（1）如果新加入的交换机本身未形成堆叠，则新加入的交换机会被选为从交换机，堆叠系统中原有主、备角色不变。

（2）如果新加入的交换机本身已经形成了堆叠，此时相当于两个堆叠合并。在这种情况下，两个堆叠系统的主交换机将选举出一个更优的交换机作为新堆叠系统的主交换机，其中一个堆叠系统（新主交换机所在堆叠系统）将保持不变，业务也不会受到影响； 而另外一个堆叠系统的所有交换机将重新启动后加入新堆叠，并将同步主交换机的配置，该堆叠的原有业务也将中断。

6．堆叠成员退出

iStack 堆叠成员退出是指成员交换机从堆叠系统中离开，断开堆叠连接。此时会因退出成员的角色不同对堆叠系统的影响有所不同。具体如下。

（1）主交换机退出：备交换机升级为主交换机，更新堆叠拓扑结构并指定一个新的备交换机。

（2）备交换机退出：主交换机更新堆叠拓扑结构并指定一个新的备交换机。

（3）从交换机退出：主交换机更新堆叠拓扑结构。

7．堆叠主、备切换和堆叠系统MAC地址切换

当iStack堆叠系统成功建立后，如果主交换机故障或脱离堆叠系统，则备交换机自动提升为主交换机，然后由新的主交换机指定新的备交换机，进行主、备交换机数据同步。这里的堆叠主、备切换，以及堆叠系统MAC地址的切换又要区分以下3种情况。

（1）当堆叠系统第一次成功建立之后，此时堆叠系统的 MAC 地址是主交换机的MAC地址。当主交换机发生故障或脱离堆叠系统时，在去使能堆叠系统MAC地址延时切换功能的情况下，系统 MAC 地址会立刻切换为新的主交换机的 MAC 地址。缺省使能堆叠系统MAC地址延时切换功能，延迟时间为10min。

（2）当堆叠系统成功建立之后，如果主交换机故障或脱离堆叠系统，如果堆叠系统配置了系统MAC地址切换时间，且在切换定时器超时时间内旧主交换机还没有重新加入堆叠系统，则新主交换机将堆叠系统的MAC地址切换为自己的MAC地址；反之，如果在切换定时器超时时间内旧主交换机重新加入堆叠，此时系统旧主交换机变为从交换机，但堆叠系统的**MAC**地址不切换 。相当于，此时堆叠系统的**MAC**地址为从交换机**MAC**地址 。

（3）当堆叠交换机中有从交换机离开时，如果离开的从交换机的 MAC 地址是堆叠的系统MAC地址（如上面这种情况），且该交换机在切换定时器超时时间内没有重新加入堆叠，则主交换机将堆叠系统MAC地址切换为自己的MAC地址。

但要注意，频繁的主备切换有可能导致堆叠分裂。

8．堆叠分裂

iStack堆叠分裂是指稳态运行的堆叠系统中带电移出部分成员交换机，或者堆叠线缆多点故障导致一个堆叠系统变成多个堆叠系统，如图5-6所示。堆叠系统分裂后，可能产生多个有相同配置的堆叠系统，导致网络中IP地址和MAC地址的冲突，引起网络故障。

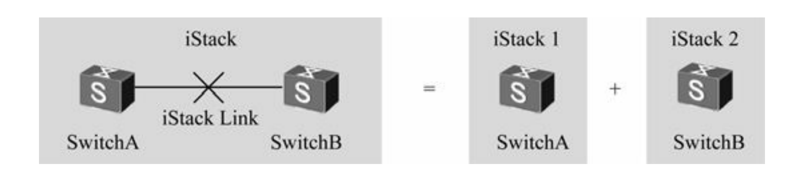


图5-6 堆叠分裂示意图

9．双主检测

双主检测DAD（Dual-Active Detection），是一种检测和处理堆叠分裂的协议，可以实现堆叠分裂的检测、冲突处理和故障恢复，降低堆叠分裂对业务的影响，仅**S5700**和**S6700**系列支持，且仅支持由两台交换机组成的堆叠系统 。

双主检测方式有两种：直连检测方式和Relay代理检测方式。

（1）直连检测方式。如图5-7所示，堆叠成员交换机间通过专用直连链路进行双主检测。在直连检测方式中，堆叠系统正常运行时，为了减轻 CPU 负担不发送 DAD 报文；堆叠系统分裂后，堆叠成员交换机以 1s 为周期通过检测链路发送DAD报文。

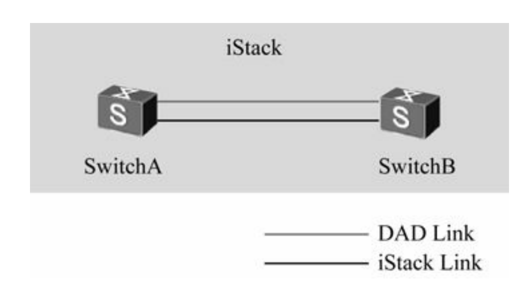


图5-7 直连方式双主检测示意图

（2）Relay代理检测方式。如图5-8所示， Relay 代理检测方式在堆叠系统跨交换机Eth-Trunk上启用DAD检测，在代理交换机上启用DAD代理功能。代理交换机必须为支持DAD Relay代理功能的交换机，目前 S系列交换机都支持DAD Relay代理功能。

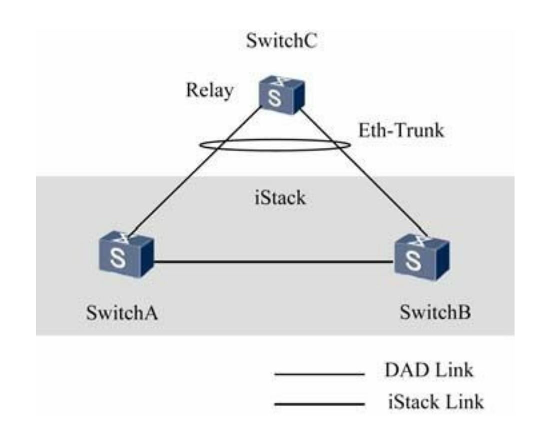


图5-8 Relay代理方式双主检测示意图

在Relay代理检测方式中，堆叠系统正常运行时堆叠成员交换机以 30s 为周期通过检测链路发送DAD报文。堆叠成员交换机对在正常工作状态下收到的DAD报文不做任何处理；堆叠 系统分裂后，堆叠成员交换机以1s为周期通过检测链路发送DAD报文。

堆叠分裂后，分裂成多部分的堆叠系统会在检测链路上相互发送 DAD 竞争报文。堆叠系统将接收到的报文信息与本部分竞争信息做比较：如果本部分竞争为主交换机则不做处理，保持Active状态，正常转发业务报文；如果本部分竞争为备交换机，则需要关闭除保留端口（交换机上不会被关闭的端口）外的所有业务端口，转入Recovery状态，停止转发业务报文。堆叠链路修复后，处于Recovery状态的堆叠将重新启动，同时将被关闭的业务端口恢复Up，整个堆叠系统恢复。

5.2 iStack配置与管理

了解了iStack基础知识后，下面正式介绍iStack堆叠的配置与管理方法。在这里同样要注意，S2700、S3700系列与S5700、S6700系列在iStack堆叠的配置上同样存在一些区别，具体将在下面介绍的配置步骤中体现。

各S系列交换机与iStack堆叠有关的参数的缺省配置：堆叠使能状态已使能，堆叠ID为0，堆叠优先级为100。

5.2.1 iStack堆叠配置任务

iStack 堆叠的配置不是很复杂，最基本的配置总的来说包括几个方面：使能堆叠功能，指定堆叠端口，配置堆叠ID和堆叠优先级，当然还可以有一些其他可选配置任务。堆叠卡连接方式和业务口连接方式的堆叠建立流程分别如图5-9左、右图所示。

重新启动交换机使堆叠建立后，用户可以根据实际需求有选择地配置堆叠系统保留VLAN、堆叠系统MAC地址的切换时间和接口指示灯显示堆叠ID。

下面综合介绍以上两种堆叠连接方式下的堆叠基本配置任务，注意这些配置任务中绝大多数是可选的，而且不同S系列交换机中在配置任务上也有所不同。

1．（可选）配置堆叠端口（业务口连接方式必选 ）

本项配置任务仅**S5700LI**、**S5710EI**和**S6700**系列交换机需要，使用的是普通的业务口作为堆叠端口的物理成员端口，以堆叠卡方式连接的 S2700、S3700、S5700SI 和S5700EI系列交换机不支持此配置。

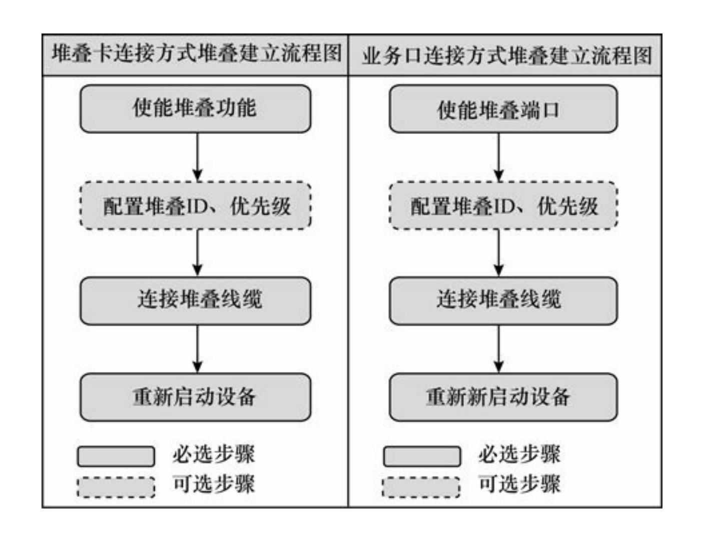


图5-9 两种堆叠连接方式下的堆叠建立示意图

当堆叠成员交换机之间通过业务口方式连接时，需要将普通的业务口配置为堆叠物理成员端口，并将其加入到逻辑堆叠端口中。堆叠端口必须和堆叠物理成员端口绑定才能有效。一个堆叠端口中可以加入多个堆叠物理成员端口，以提高堆叠链路的带宽和可靠性，具体能绑定多少个物理成员端口，不同系列交换机也不一样，参见本章表5-1说明。

2．（可选）使能堆叠功能（堆叠卡连接方式必选 ）

本项配置任务仅**S5700EI**和**S5700SI**子系列，以及**S2700**和**S3700**系列交换机需要，使用的是专门的堆叠卡中的端口作为堆叠端口，S5700LI、S5710EI和S6700系列交换机中以业务口方式连接的堆叠不支持此配置。

3．（可选）配置堆叠ID

堆叠ID用来标识和管理堆叠中的成员交换机，堆叠中所有成员交换机的堆叠ID都是唯一的。堆叠系统建立时，如果成员交换机的堆叠 ID 有冲突，主交换机将为冲突的成员交换机重新分配堆叠ID，也以手工指定。

4．（可选）配置堆叠优先级

堆叠优先级主要用于角色选举过程中确定成员交换机的角色，优先级值越大表示优先级越高，优先级

越高当选为主交换机的可能性越大。每台成员交换机的缺省优先级都是100，也可手工指定，最好指定为不同优先级。

5. 重新启动交换机

完成以上基本配置后，使用专用的堆叠线缆连接两个堆叠端口，然后需要重新启动 交换机，使以上基本配置生效，以建立堆叠。重新启动交换机使堆叠建立后，用户可以根据实际需求有选择地进行下面的配置任务，如配置堆叠系统保留 VLAN、堆叠系统MAC地址的切换时间和接口指示灯显示堆叠ID等。为了成功组建堆叠，用户可以通过命令行或手动重新启动交换机。

6. （可选）配置主备倒换

在iStack堆叠系统建立之后，如果主交换机出现故障或者已脱离堆叠系统，则原来的备交换机会自动提升为主交换机，然后根据选举规则确定新的备交换机，并进行主备数据同步。这是自动切换的情况，还有一种人工强制倒换主、备交换机角色的情况。

在自动切换和手工强制倒换的这两种情况下都涉及一个堆叠系统 MAC 地址的重新确定问题，具体要依据以下3种情况而定（当堆叠系统第一次成功建立之后，此时堆叠系统的MAC地址是主交换机的MAC地址）。

（1）如果去使能了堆叠系统MAC地址延时切换功能，则堆叠系统MAC地址会立刻切换为新主交换机的MAC地址（缺省使能堆叠系统MAC地址延时切换功能，延迟10min切换）。

（2）如果使能了堆叠系统 MAC 地址延时切换功能，并配置了堆叠系统 MAC 地址切换时间，则在切换定时器超时时间内，如果旧主交换机还没有重新加入堆叠，新主交换机将堆叠系统的MAC地址切换为自己的MAC地址；如果在切换定时器超时时间内，旧主交换机重新加入了堆叠，则旧主交换机变为从交换机角色，但堆叠系统的 MAC 地址不切换，为旧主交换机（现为从交换机）的MAC地址。

（3）如果现为从交换机，并且其 MAC 地址是堆叠的系统 MAC 地址的原旧主交换机再次脱离堆叠系统，则当该交换机在切换定时器超时时间内没有重新加入堆叠时，新主交换机将堆叠系统MAC地址切换为自己的MAC地址。

7. （可选）配置堆叠系统保留VLAN

缺省情况下，堆叠系统使用VLAN 4093作为堆叠系统的保留VLAN，用于堆叠协议报文的交互，其他业务不能使用此VLAN。如果用户需要使用VLAN 4093来部署业务，可以通过此命令来修改堆叠系统的保留VLAN。

8. （可选）配置堆叠系统MAC地址的切换时间

当堆叠交换机中有成员交换机离开，且如果离开的成员交换机的 MAC 地址是原来堆叠系统的MAC地址。为了在这种情况下不至于出现堆叠MAC地址无法确定的现象，可在堆叠系统配置MAC地址的切换时间，使得MAC地址为堆叠系统MAC地址的某交换机离开堆叠系统，且在指定的时间没有重新加入堆叠系统时，则主交换机将堆叠系统MAC地址切换为自己的MAC地址。但这里要区分两种情况。

（1）如果在堆叠系统建立之前在所有成员交换机上已各自完成了系统MAC地址切换时间的配置，那么在堆叠系统建立之后系统MAC地址切换时间为主交换机上配置的系统MAC地址切换时间，并且其他成员交换机的系统MAC地址切换时间和主交换机上的系统MAC地址切换时间配置保持一致。当堆叠系统重启出现了主、备倒换，则堆叠系统MAC地址切换为新主交换机的MAC地址；当主交换机不变时系统MAC地址不发生改变，不会发生MAC地址切换。

（2）如果是在堆叠系统建立之后配置系统MAC地址切换时间，则在发生主、备倒换后，但原主交换机仍在堆叠系统中时，堆叠系统重启后系统MAC地址不发生改变， 不会发生MAC地址切换。

9. （可选）配置接口指示灯显示堆叠ID

堆叠交换机的堆叠ID可以在堆叠建立时由主交换机自由分配，也可以由用户设定。如果由主交换机自由分配，具体哪台交换机对应的 ID 就很难分辨，此时若想对堆叠中的具体一台交换机进行操作将无法确定其所在位置。这时可以通过配置接口指示灯指示对应的堆叠ID就可直观地显示交换机的堆叠ID了。在华为S系列交换机上通过接口指示灯判断堆叠ID的规则如下。

（1）堆叠ID为1～8号的交换机：只有与堆叠ID对应的序号端口灯亮。例如，堆叠ID为1，则第一个端口灯亮，堆叠ID为2，则第二个端口灯亮。

（2）堆叠ID为0号的交换机：本产品支持几台交换机堆叠，就是前面几个端口灯全亮。例如，某产品支持9台交换机堆叠，则其前面9个端口灯全亮表示本交换机堆叠ID为0。

配置接口指示灯显示堆叠 ID 后，主交换机的接口指示灯亮灯状态为闪烁，而从交换机的接口指示灯状态为常亮。

5.2.2 配置iStack堆叠

本节是依据上节介绍的配置任务来介绍各项配置任务的具体配置方法，如表5-2所示（如果不采用缺省配置，则需要在每台成员交换机上完成其中的配置 ）。但要区分两种不同的堆叠连接方式和不同机型对相应配置步骤的支持。

表5-2 iStack堆叠的配置步骤

<table>
<tr><th>配置任务</th><th>步骤</th><th>命令</th><th colspan="2">说明</th></tr>
<tr><td rowspan="3">创建链路聚合组（仅适用于采用业务口连接方式的 S 系列交换机，参见表 5-1）</td><td>1</td><td>system-view
例如：<HUAWEI> system-view</td><td colspan="2">进入系统视图</td></tr>
<tr><td rowspan="2">2</td><td>stack port interface interface-type interface-number enable
例如：[HUAWEI]stack port interface gigabitethernet 0/0/28 enable</td><td rowspan="2">（二选一）配置业务口为堆叠物理成员端口，根据机型选择其中一个命令</td><td>仅适用于 S5700LI 和 S5710EI 子系列，只有最后 4 个业务口可以配置为堆叠物理成员端口，但 S5710EI 子系列还可通过子卡扩展 4 个堆叠物理成员端口。可用于配置物理成员端口的子卡为 ES5D21X02S00
缺省情况下，业务接口未配置为堆叠物理成员端口，可用 undo stack port interface interface-type interface-number enable 命令恢复堆叠物理成员端口为业务口</td></tr>
<tr><td>stack port interface interface-type interface-number1 to interface-number2 enable
例如：[HUAWEI] stack port interface gigabitethernet 0/0/5 to 0/0/8 enable</td><td>仅适用于 S6700 系列，最多可以配置 8 个业务口为堆叠物理成员端口（端口号不固定），但必须为 4 的倍数个连续的一组接口同时配置，例如可以是 1～4、5～8 或者 1～8，但不可以是 2～5、3～6
缺省情况下，业务接口未配置为堆叠物理成员端口，可用 undo stack port interface interface-type interface-number1 to interface-number2 enable 命令恢复堆叠物理成员端口为业务口</td></tr>
</table>

（续表）

配置任务	步骤	命令	说明
创建链路聚合组（**仅适用于采用业务口连接方式的 S 系列交换机，参见表 5-1**）	3	**interface stack-port** *member-id/port-id* 例如：[HUAWEI] **interface stack-port** 1/1	进入逻辑堆叠端口视图，**仅 S5700LI、S5710EI 和 S6700 系列支持此命令**。参数对 *member-id/ port-id* 用来分别指定堆叠成员交换机的堆叠 ID 和堆叠端口编号，取值范围分别为 0～8 和 0～2 的整数 iStack 堆叠的每台成员交换机上有两个堆叠端口，即 Stack-Port*n*/1 和 Stack-Port*n*/2，其中 *n* 为成员交换机的堆叠 ID。执行本命令进入指定的堆叠端口视图后，可以配置对应成员交换机的堆叠端口的相关属性
	4	**port member-group interface** *interface-type interface-number* 例如：[HUAWEI-stack-port0/1]**port member-group interface** gigabitethernet **0/0/28**	向以上逻辑堆叠端口中添加堆叠物理成员端口，**仅 S5700LI、S5710EI 和 S6700 系列支持此命令** 【注意】在向堆叠端口中添加成员端口以及在堆叠端口连接时要注意以下事项。 （1）堆叠成员交换机之间本端交换机的堆叠端口 1 必须与对端交换机的堆叠端口 2 相连 （2）S5700LI 子系列每个堆叠端口最多只能添加 2 个堆叠物理成员端口，可以配置的 4 个堆叠物理成员端口中 **GigabitEthernet0/0/25 和 GigabitEthernet0/0/26 必须加入同一个堆叠端口，GigabitEthernet0/0/27 和 GigabitEthernet0/0/28 必须同时加入另一个堆叠端口** （3）S5710EI 子系列最多只能添加 4 个堆叠物理成员端口，**但子卡上的接口和交换机面板上的接口不能混合加入同一个堆叠端口，不同子卡上的接口可加入同一个堆叠端口** （4）S6700 系列最多只能添加 8 个堆叠物理成员端口，**且加入方式不受限制** 缺省情况下，堆叠端口中没有堆叠物理成员端口，可用 **undo port member-group interface** *interface-type interface-number* 命令删除堆叠端口中指定的堆叠物理成员端口
使能堆叠功能	5	**quit** 例如：[HUAWEI-stack-port0/1]**quit**	退出堆叠端口视图，返回系统视图
	6	**stack enable** 例如：[HUAWEI] **stack enable**	使能交换机堆叠功能，**仅适用采用堆叠卡连接方式的 S5700EI、S5700SI 子系列交换机和 S2700、S3700 系列交换机**。交换机在建立堆叠系统之前必须要使能堆叠功能 缺省情况下，交换机的堆叠功能处于使能状态，可用 **undo stack enable** 命令去使能交换机的堆叠功能
（可选）配置交换机堆叠 ID	7	**stack slot** *slot-id* **renumber** *new-slot-id* 例如：[HUAWEI] **stack slot** 4 **renumber** 5	指定成员交换机的堆叠 ID。命令中的参数说明如下。 （1）*slot-id*：指定需要修改堆叠 ID 的成员交换机堆叠 ID，取值范围为 0～8 的整数，缺省堆叠 ID 为 0 （2）*new-slot-id*：指定修改后的堆叠 ID，取值范围为 0～8 的整数 修改交换机的堆叠 ID 时会有确认提示，确认后才会执行。**配置交换机的堆叠 ID 后，需要重启交换机配置才能生效，但执行该命令前必须先使能成员交换机的堆叠功能**

（续表）

配置任务	步骤	命令	说明
（可选）配置交换机堆叠优先级	8	**stack slot** *slot-id* **priority** *priority* 例如：[HUAWEI] **stack slot** 5 **priority** 150	配置成员交换机的堆叠优先级。命令中的参数说明如下。 （1）*slot-id*：指定要修改堆叠优先级的成员交换机堆叠 ID，取值范围为 0～8 的整数 （2）*priority*：指定修改后的堆叠优先级，取值范围为 1～255 的整数，缺省堆叠优先级为 100。值越大，优先级越高，交换机被选为主交换机的可能性越大 修改交换机的堆叠 ID 时会有确认提示，确认后才会执行，并且**执行该命令前必须保证交换机处于堆叠状态**
重新启动交换机	9	**reboot** 例如：[HUAWEI] **reboot**	重新启动交换机。交换机配置堆叠相关属性后，如修改了堆叠 ID 和使能了堆叠功能，并使用了专用的堆叠线缆连接两个堆叠端口，需要重新启动交换机后配置才能生效。为了成功组建堆叠，用户可以通过本命令或手动重新启动交换机。但重启交换机前使用 **save** 命令保存配置
（可选）配置主备倒换	10	**display switchover state** 例如：[HUAWEI]**display switchover state**	查看系统是否满足主备倒换的条件，要保证备交换机处于实时备份阶段（输出信息显示为"receiving realtime or routine data"）
	11	**slave auto-update config** 例如：[HUAWEI] **slave auto-update config**	使能主、备交换机配置数据同步功能。使能主备交换机配置数据同步功能后，如果在主交换机上执行命令 **save** 保存配置，系统自动将配置信息同步到备交换机，实现主、备交换机上的配置文件同步 缺省情况下，主备交换机配置数据自动同步，可用 **undo slave auto-update config** 命令去使能主备交换机配置数据同步功能
	12	**slave switchover enable** 例如：[HUAWEI]**slave switchover enable**	使能强制进行主备倒换功能。缺省情况下，主备倒换功能处于使能状态，可用 **slave switchover disable** 命令禁止强制进行主备倒换功能
	13	**slave switchover** 例如：[HUAWEI] **slave switchover**	强制进行主备倒换。只有在使能强制进行主、备倒换功能之后，用户才可以使用本命令进行主备倒换；如果禁止强制进行主备倒换功能，则用户无法通过配置命令强制进行主、备倒换
（可选）配置堆叠系统保留 VLAN	14	**stack reserved-vlan** *vlan-id* 例如：[HUAWEI]**stack reserved-vlan** 2000	配置堆叠系统保留 VLAN，参数 *vlan-id* 用来指定要保留给堆叠系统使用的 VLAN ID，取值范围为 1～4 094 的整数 缺省情况下，堆叠系统使用 VLAN 4093 作为堆叠系统的保留 VLAN，用于堆叠协议报文的交互，其他业务就不能使用此 VLAN 如果需要使用 VLAN 4093 来部署业务，可以通过此命令来修改堆叠系统的保留 VLAN。由于使用保留 vlan 部署业务时会导致堆叠无法组建，因此新指定堆叠系统的保留 VLAN 必须是没有部署于其他业务的

（续表）

配置任务	步骤	命令	说明
（可选）配置堆叠系统 MAC 地址的切换时间	15	**stack timer mac-address switch-delay** *delay-time* 例如：[HUAWEI] **stack timer mac-address switch-delay** 5	配置系统 MAC 地址的切换时间，参数的 *delay-time* 用来指定系统 MAC 地址的切换时间，取值范围为 0～60min 当堆叠交换机中有成员离开，如果离开的交换机的 MAC 地址是堆叠的系统 MAC 地址，则该交换机在参数 *delay-time* 指定的时间内没有重新加入堆叠，主交换机将堆叠系统 MAC 地址切换为自己的 MAC 地址。堆叠中的所有成员交换机的系统 MAC 地址切换时间与主交换机一致。当参数 *delay-time* 的值设置为 0 时，表示不切换 缺省情况下，系统 MAC 地址的切换时间为 10min，可用 **undo stack timer mac-address switch-delay** 命令恢复为缺省值
（可选）配置接口指示灯显示堆叠 ID	16	**stack led enable** [**duration** *duration-value*] 例如：[HUAWEI] **stack led enable duration** 60	使能接口指示灯显示堆叠交换机的堆叠 ID 功能。命令中的可选参数 *duration-value* 指定每次接口指示灯显示堆叠 ID 的持续时间，取值范围为 30～600 的整数秒，缺省为 45s 缺省情况下，交换机未使能接口指示灯显示堆叠交换机的堆叠 ID 功能，可用 **stack led disable** 命令去使能该功能

【示例 1】将堆叠ID为4的交换机的堆叠ID修改为5。

< HUAWEI > system-view

[HUAWEI] stack slot 4 renumber 5

Warning:Please do not frequently modify slotid, it will make the stack split! co

ntinue?[Y/N]:y

Info: Stack configuration has been changed, need reboot to take effect.

【示例 2】将本交换机的系统MAC地址切换时间配置为4min。

< HUAWEI > system-view

[HUAWEI] stack timer mac-address switch-delay 4

Warning:Please do not frequently modify mac switch time, it will make the stack

split! continue?[Y/N]:y

【示例 3】查看系统是否满足主备倒换的条件（仅当备交换机处于实时备份状态才满足条件）。输出信息字段说明如表5-3所示。

< HUAWEI > display switchover state

Slot 1 HA FSM State(master): waiting for the slave to be inserted.

表5-3 display switchover state命令输出信息字段说明

字段	说明
HA FSM State(master)	表示主交换机的备份状态，有以下几种状态。 （1）waiting for the slave to be inserted：表示目前只有主交换机而没有备交换机 （2）waiting batch backup request from slave：表示主交换机在等待备交换机的批量备份请求 （3）realtime or routine backup：表示主交换机处于实时备份状态 （4）data smooth：表示主交换机处于数据平滑状态
HA FSM State(slave)	表示备交换机的备份状态，有以下几种状态。 （1）ready：表示备交换机已经完成启动，并准备接收批量备份数据 （2）receiving batch data：表示备交换机正在接收批量备份数据 （3）receiving realtime or routine data：表示备交换机处于实时接收数据的状态（**仅此状态满足主备倒换条件**）

【示例 4】配置系统进行主备倒换。倒换前会有系统确认提示，只有键入 Y 或者 y才执行主、备倒换操作。

< HUAWEI > system-view

[HUAWEI] slave switchover enable

[HUAWEI] slave switchover

Warning: This operation will switch the slave board to the master board. Continue? [Y/N]:y

5.2.3 iStack堆叠管理

在配置好上节介绍的iStack堆叠后以及iStack堆叠使用过程中，可在任意视图下执行以下display命令查看相关配置和统计信息，使用以下reset用户视图命令清除iStack统计信息。

（1）使用display stack命令查看堆叠成员交换机堆叠信息，包括堆叠拓扑和堆叠成员等信息。

（2）使用display stack peers命令可查看堆叠系统中成员交换机各堆叠端口相连的邻居信息。

（3）使用 display stack configuration [slot slot-id] 命令可查看所有或者指定堆叠ID的成员交换机的堆叠配置信息，包括堆叠成员交换机当前以及下次启动时的堆叠ID、优先级信息。可选参数slot-id用来指定要查看堆叠配置信息的成员交换机的堆叠ID，取值范围为0～8的整数。

（4）在 S5700LI和 S5710EI子系列交换机中使用 display stack-port membership [slot-i d/port-id] 命令查看指定或所有成员交换机/堆叠端口下的成员端口信息。

（5）在 S6700系列中使用 display stack-port { global load-balance | load-balance [slot-id/port-id] |membership [slot-id/port-id] }命令查看所有成员交换机的负载分担模式，或者指定成员交换机/堆叠端口下的成员端口信息。

（6）使用 display interface stack-port [slot-id/port-id] 命令可查看所有或者指定堆叠成员交换机/堆叠端口下各成员接口的流量统计信息。可选参数slot-id/port-id用来指定要查看流量统计信息的堆叠ID/堆叠端口。

（7）使用 reset counters stack-port [slot-id/port-id] 命令清除所有或指定成员交换机/堆叠端口的状态统计信息。如果不指定堆叠成员交换机的堆叠 ID/堆叠端口编号，则清除所有类型端口的统计信息；如果指定堆叠成员交换机的堆叠 ID/堆叠端口编号，则清除指定端口的统计信息。

5.2.4 iStack堆叠配置示例

本示例拓扑结构如图5-10所示，SwitchA、SwitchB、SwitchC和SwitchD四台交换机组成环型堆叠系统，以满足网络规模的扩大，以及日益增加的接入层交换机提供的端口数要求。本示例仅以支持普通业务口连接方式的S5700LI子系列交换机为例进行介绍。

在S5700LI子系列交换机中是以最后2个（每个堆叠端口一个成员端口时）或4个（每个堆叠端口两个成员端口时）业务端口作为堆叠端口的成员端口的，所以需要用SFP+堆叠电缆连接各成员端口，但要注意的是同一条堆叠链路上的两个堆叠物理成员端口需加入不同堆叠端口，即本端交换机的堆叠端口**1**必须与对端交换机的堆叠端口**2**相连 。

因为本示例只是最基本的堆叠组建，又是采用业务口堆叠连接方式，所以根据5.2.2节介绍的具体配置步骤可以很容易得出本示例只需要三项主要配置任务（其他可选配置任务均采用缺省配置），即指定物理业务端口作为堆叠端口的成员端口，配置堆叠成员ID和堆叠优先级，重启各成员交换机。在进行正式配置前需按照图 5-10 所示拓扑结构使用堆叠电缆连接好各成员交换机。以下是这三项配置任务的具体配置步骤。

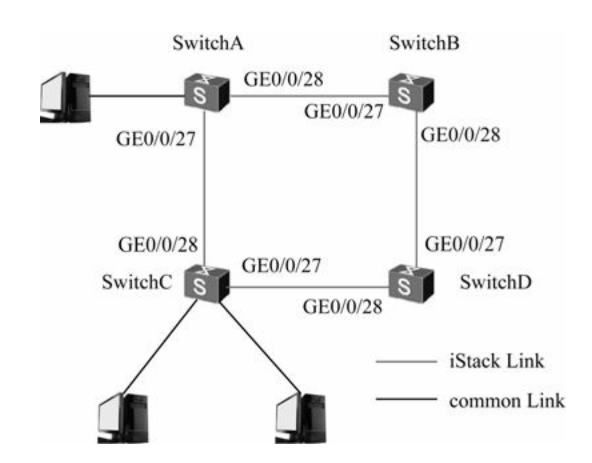


图5-10 S5700LI子系列交换机堆叠配置示例拓扑结构

（1）在各成员交换机上配置堆叠端口，添加物理成员业务端口。因为各成员交换 机是同型号的，而且在配置上也完全一样，所以下面仅以SwitchA上的配置为例进行介绍，其他成员交换机的配置完全一样，参见即可。

将SwitchA的最后两个业务端口GigabitEthernet0/0/27～GigabitEthernet0/0/28分别配置为1号和2号堆叠端口的物理成员端口，并加入堆叠端口。注意此时各成员交换机的堆叠ID均为0，等下一步再来修改各交换机的堆叠ID。

```
<HUAWEI>system-view
[HUAWEI] sysname SwitchA
[SwitchA] stack port interfacegigabitethernet0/0/27 enable
Warning: Enabling stack port may cause configuration loss on the interface, cont
inue?[Y/N]:y
[SwitchA] stack port interfacegigabitethernet0/0/28 enable
Warning: Enabling stack port may cause configuration loss on the interface, cont
inue?[Y/N]:y
[SwitchA] interface stack-port 0/1
[SwitchA-stack-port0/1] port member-group interfacegigabitethernet0/0/27
[SwitchA-stack-port0/1] quit
[SwitchA] interface stack-port 0/2
[SwitchA-stack-port0/2] port member-group interfacegigabitethernet0/0/28
[SwitchA-stack-port0/2] quit
```

（2）配置各成员交换机的堆叠ID和堆叠优先级。现在假设以SwitchA为主交换机，现只要为SwitchA配置优先级（假设为200），其他成员交换机的优先级保持缺省值100即可使SwitchA成为主交换机。在堆叠ID方面，SwitchA采用缺省的0号，SwitchB、SwitchC和SwitchD分配的堆叠ID分别为1、2、3。

下面仅介绍SwitchA的优先级配置（其他交换机的优先级不用配置，直接采用优先级较低的缺省值**100**）和SwitchB的堆叠ID（SwitchA的堆叠ID不用配置，SwitchB、SwitchC和SwitchD的堆叠ID配置方法与SwitchB的一样）配置方法。

配置SwitchA的堆叠优先级为200。

```
[SwitchA] stack slot 0 priority200
```

Warning:Please do not frequently modify Priority, it will make the stack split!

continue?[Y/N]:y

配置SwitchB的堆叠ID为1。

[SwitchB] stack slot 0 renumber1

Warning:Please do not frequently modify slotid, it will make the stack split! co

ntinue?[Y/N]:y

Info: Stack configuration has been changed, need reboot to take effect.

（3）在各交换机上执行 save 用户视图命令保存配置在配置文件中，使用 reboot 用户视图命令重启各成员交换机。

（4）重启各成员交换机后在堆叠主交换机SwitchA的任意视图下执行display stack命令检查以上配置结果，验证是否正确。下面仅以SwitchA为例进行介绍，输出信息如下：

```
<SwitchA>display stack
Stack topology type : Ring
Stack system MAC：0018-82d2-2e85
MAC switch delay time: 10 min
Stack reserve vlanid : 4093
slot#   Role      Mac address      Priority    Device type
-------------------------------------------------------------
  0     Master    0018-82d2-2e85   200         S5700-28P-LI-AC
  1     Slave     0018-82c6-1f44   100         S5700-28P-LI-AC
  2     Standby   0018-82c6-1f4c   100         S5700-28P-LI-AC
  3     Slave     0018-82b1-6eb8   100         S5700-28P-LI-AC
```

5.2.5 双主检测配置与管理

通过配置堆叠双主检测，可以检测并处理堆叠分裂时网络中出现的双主冲突，但仅支持由两台交换机组成的堆叠系统进行双主检测。本特性仅在**S5700**和**S6700**系列中支持**iStack**堆叠的子系列交换机支持 。堆叠双主检测方式有两种。

（1）直连检测方式：成员交换机间通过专用直连链路进行双主检测。

（2）Relay代理检测方式：成员交换机间通过启用Relay代理功能的交换机进行双主检测。

以上两种检测方式的拓扑结构分别参见5.1.2节的图5-7和图5-8。

说明

DAD（双主检测）为华为公司的私有协议，华为公司生产的所有S系列交换机都 支持DAD代理功能。DAD报文是BPDU报文，在代理交换机和直连检测链路的中间交换机上需要配置接口转发BPDU报文。缺省情况下，接口上送BPDU报文到CPU处理。

在同一个堆叠系统中，两种检测方式互斥，不可以同时配置。为了保证检测的可靠性，可以同时配置4条直连检测链路或4个Relay代理检测Eth-Trunk，在出现双主时，只要有1条直连检测链路或1个Relay代理处于Up状态即可确保工作正常。

1．配置任务

双主检测的配置任务包括以下3项（仅第一项是必选的 ）：

（1）配置检测方式：可以是直连检测方式或者Relay代理检测方式。

（2）（可选）配置保留端口。双主检测发现堆叠分裂故障后，为防止相同的MAC地址、IP 地址引起网络振荡，需要将竞选失败的成员交换机上的所有业务端口关闭，以减少对网络的影响。如果有部分端口仅做报文透传功能，出现双主故障时不会影响到网络运行，这时可以通过命令将这些端口配置为保留端口，在双主故障时关闭除保留端口外的所有业务端口。

（3）（可选）恢复被关闭的端口。如果在堆叠系统分裂故障恢复前，原主交换机发生故障或被移出网络，则可以通过命令行先启用处于端口关闭状态的备份交换机，使所有业务口重新恢复正常，让它接替原主交换机工作，以保证业务尽量少受影响。

2．配置步骤

以上三项双主检测具体配置任务的配置步骤如表5-4所示。

表5-4 双主检测的配置步骤

配置任务	步骤	命令		说明	
配置检测方式	1	**system-view** 例如：<HUAWEI> **system-view**		（二选一）配置直连检测方式	进入系统视图
	2	**interface** *interface-type interface-number* 例如：[HUAWEI]**interface** gigabitethernet 1/0/1			键入直连用于双主检测的以太网端口，进入接口视图
	3	**mad detect mode direct** 例如：[HUAWEI-GigabitEthernet1/0/1]**mad detect mode direct**			配置以上接口的直连双主检测功能。缺省情况下，接口的直连双主检测功能处于关闭状态 【说明】配置指定接口的直连方式双主检测功能后，该接口会进入阻塞状态。取消配置指定接口采用直连方式双主检测功能后，接口会恢复转发功能，所以在接口被配置直连双主检测功能后，不要再配置其他业务 缺省情况下，接口的直连双主检测功能处于关闭状态，可用 **undo mad detect mode direct** 命令去使能该功能
	2	**interface eth-trunk** *trunk-id* 例如：[HUAWEI] **interface eth-trunk** 2	在堆叠交换机上配置Relay代理检测功能	（二选一）配置Relay代理检测方式	进入堆叠交换机的 Eth-Trunk 接口视图
	3	**mad detect mode relay** 例如：[HUAWEI-Eth-Trunk2]**mad detect mode relay**			在以上堆叠交换机 Eth-Trunk 接口上使能 Relay 代理双主检测功能。缺省情况下，未使能 Relay 代理双主检测功能，可用 **undo mad detect mode relay** 命令去使能该功能
	4	**interface eth-trunk** *trunk-id* 例如：[HUAWEI] **interface eth-trunk** 2	在代理交换机上配置Relay代理检测功能		进入代理交换机的 Eth-Trunk 接口视图
	5	**mad relay** 例如：[HUAWEI-Eth-Trunk2]**mad relay**			在以上代理交换机 Eth-Trunk 接口上使能 Relay 代理功能 缺省情况下，接口未使能 Relay 代理功能，可用 **undo mad relay** 命令去使能该功能
（可选）配置保留端口	6	**mad exclude interface** { *interface-type interface-number1* [**to** *interface-type interface-number2*] } &<1-10> 例如：[HUAWEI] **mad exclude interface** gigabitethernet 1/0/2 **to** gigabitethernet 1/0/3		配置堆叠交换机中指定的端口为保留端口。缺省情况下，堆叠物理成员端口为保留端口，其他所有业务口均为非保留端口，可用 **undo mad exclude interface** { *interface-type interface-number1* [**to** *interface-type interface-number2*] } &<1－10>命令取消堆叠系统指定接口为保留端口 【说明】如果有部分端口仅做报文透传功能，出现双主故障时不会影响到网络运行，这时可以通过本命令将这些端口配置为保留端口，在出现双主故障时将关闭除保留端口外的所有业务端口	

（续表）

配置任务	步骤	命令	说明
（可选）恢复被关闭的端口	7	**mad restore** 例如：[HUAWEI]**mad restore**	将堆叠分裂后进入 Recovery 状态的设备被关闭的所有端口重新恢复正常 【说明】双主检测发现堆叠分裂产生的双主冲突后，竞选成功的成员交换机保持为 Active 状态，竞选失败的成员交换机除保留端口外的所有业务口关闭，转入 Recovery 状态，暂时不能转发业务报文。如果分裂故障还没来得及修复而处于 Active 的交换机也故障了或被移出了网络，此时可在处于 Recovery 状态的交换机上执行本命令使该交换机上处于关闭状态的端口重新恢复正常，先接替原 Active 交换机的工作，再修复原 Active 交换机故障及堆叠链路故障。故障修复后，两台交换机重新建立堆叠系统。**但在主交换机正常工作时，不能执行该命令**，否则会再次发现双主并关闭业务端口，从而导致端口震荡

配置好后，可用 display mad verbose [proxy | verbose] 命令查看双主检测配置信息。如果选择 proxy 二选一可选项，则显示的是代理交换机的 Relay 代理信息，在用户配置了Relay代理方式的双主检测时选用；如果选择verbose二选一可选项，则显示的是堆叠双主检测详细配置信息；如果这两个可选项都不选择，则显示双主检测的摘要配置信息。

5.2.6 直连检测方式的DAD配置示例

本示例拓扑结构如图 5-11 所示，SwitchA 和SwitchB组成堆叠系统，SwitchA的堆叠ID为0， SwitchB 的堆叠 ID 为 1。在 GigabitEthernet0/0/5和 GigabitEthernet1/0/5 接口上配置采用直连双主检测功能，以减少堆叠分裂给网络带来的影响。

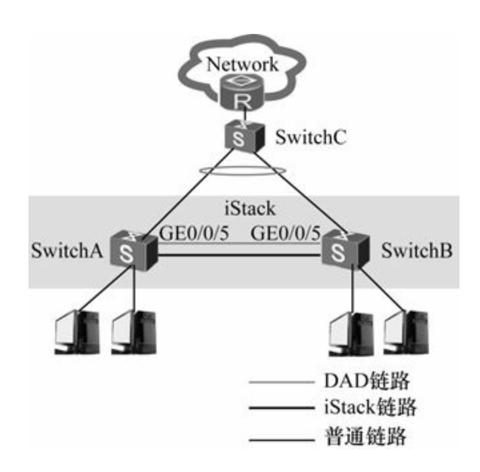


图5-11 直连检测方式的DAD配置示例拓扑结构

根据上节介绍的配置任务可以很容易得出本示例的具体配置步骤，仅需要在堆叠双方交换机用于直连检测的端口上启用直连检测功能即可。

（1）在SwitchA上配置接口GigabitEthernet0/0/5采用直连检测方式的DAD功能。

```
<HUAWEI>system-view
[HUAWEI] interface gigabitethernet 0/0/5
[HUAWEI-GigabitEthernet0/0/5] mad detect mode direct
Warning: This command will block the port, and no other configuration running on
  this port is recommended. Continue?[Y/N]:y
```

（2）在SwitchB上配置接口GigabitEthernet1/0/5采用直连检测方式的DAD功能。

<HUAWEI>system-view

[HUAWEI] interface gigabitethernet 1/0/5

[HUAWEI-GigabitEthernet1/0/5] mad detect mode direct

Warning: This command will block the port, and no other configuration running on
  this port is recommended. Continue?[Y/N]:y

配置好后，可在任意视图下通过display mad verbose命令查看堆叠系统双主检测详细配置信息，验证配置结果。

<HUAWEI>display mad verbose

Current DAD status: Detect

Mad direct detect interfaces configured:

  GigabitEthernet0/0/5

  GigabitEthernet1/0/5

Mad relay detect interfaces configured:

Excluded ports(configurable):

Excluded ports(can not be configured):

  GigabitEthernet0/0/27

  GigabitEthernet1/0/27

5.2.7 Relay代理检测方式的DAD配置示例

本示例拓扑结构如图 5-12 所示，SwitchA 和 SwitchB 组成堆叠系统，SwitchA和SwitchB通过接口Eth-Trunk1与代理交换机SwitchC连接。现配置Relay代理双 主检测功能，以减少堆叠分裂给网络带来的影响。有关 Eth-Trunk口的配置参见本书第 4章相关内容。

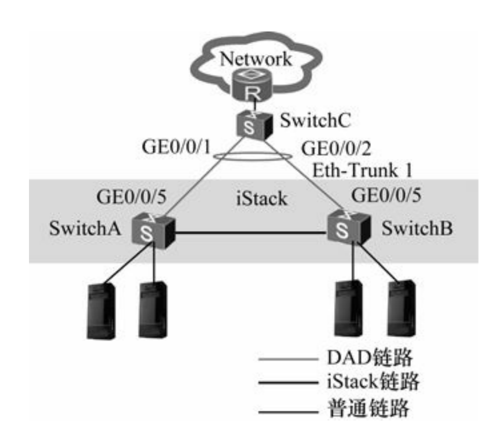


图5-12 Relay代理检测方式的DAD配置示例拓扑结构

同样根据5.2.5节介绍的DAD配置任务可以得出本示例的具体配置任务分别在堆叠交换机和代理交换机上配置用于双主检测的 Eth-Trunk链路（要向其中添加物理端口成员），并在它们的 Eth-Trunk 接口上启用Relay 代理双主检测功能。具体如下。

（1）在交换机堆叠与代理交换机SwitchC相连的Eth-Trunk接口上配置采用Relay代理检测方式的DAD功能。

```
<HUAWEI>system-view
[HUAWEI] interface eth-trunk 1
[HUAWEI-Eth-Trunk1] mad detect mode relay
[HUAWEI-Eth-Trunk1] quit
[HUAWEI] interfacegigabitethernet 0/0/5
[HUAWEI-GigabitEthernet0/0/5] eth-trunk 1
[HUAWEI-GigabitEthernet0/0/5] quit
[HUAWEI] interfacegigabitethernet 1/0/5
[HUAWEI-GigabitEthernet1/0/5] eth-trunk 1
[HUAWEI-GigabitEthernet1/0/5] quit
```

（2）在代理交换机SwitchC的Eth-Trunk接口上配置采用Relay代理检测方式的DAD功能。

```
<HUAWEI>system-view
[HUAWEI] sysname SwitchC
[SwitchC] interface eth-trunk1
[SwitchC-Eth-Trunk1] mad relay
[SwitchC-Eth-Trunk1] quit
[SwitchC] interfacegigabitethernet 0/0/1
[SwitchC-GigabitEthernet0/0/1] eth-trunk 1
[SwitchC-GigabitEthernet0/0/1] quit
[SwitchC] interface gigabitethernet 0/0/2
[SwitchC-GigabitEthernet0/0/2] eth-trunk 1
[SwitchC-GigabitEthernet0/0/2] quit
```

配置好后，同样可通过display mad verbose命令查看堆叠系统DAD配置信息，验证配置结果。具体如下。

```
<HUAWEI>display mad verbose
Current DAD status: Detect
Mad direct detect interfaces configured:
Mad relay detect interfaces configured:
  Eth-Trunk1
Excluded ports(configurable):
Excluded ports(can not be configured):
  GigabitEthernet0/0/27
  GigabitEthernet1/0/27
```

可在代理交换机SwitchC上通过display mad proxy命令查看代理信息。

```
<SwitchC>display mad proxy
Mad relay interfaces configured:
  Eth-Trunk1
```

5.3 CSS基础

随着数据中心数据访问量的逐渐增大以及网络可靠性要求越来越高，单台交换机已经无法满足需求，而通过交换机的集群能够实现数据中心大数据量转发和网络高可靠性。在华为S系列交换机中，集群技术称为CSS（Cluster Switch System，集群交换系统）。与其他交换机集群技术一样，它也是将多台支持集群特性的交换机组合在一起，从逻辑上组合成一台整体交换机。

交换机集群技术一般仅应用于高端交换机系统，如华为CSS集群技术在Sx700大系列中，只有S7700和S9700系列路由交换机支持（以前的S9300、S9300E系列也支持），主要用于提高单台交换机的转发性能和可靠性。

本章前面介绍的，在中低端S2700、S3700、S5700和S6700系列交换机中支持的iStack堆叠技术最重要的一种应用就是扩展端口，虽然也可提高单台交换机的转发性能和可靠性。这主要是因为这两种不同档次的交换机工作的层次和主要用途不一样，中低档次的交换机主要工作在接入层和汇聚层，要连接大量的终端交换机和其他网络交换机，所以需要大量的端口，而高端交换机主要应用于核心层，更需要交换机转发性能和可靠性的提高。

不过，大家往后学习就会发现，此处介绍的CSS集群技术与本章前面介绍的iStack 堆叠技术无论是在特性上，还是在配置上都存在许多相似之处，如都有主、备交换机角色，都有成员 ID 和优先级，都可以有模块卡端口连接方式和普通业务端口连接方式两种，都提供了直连检测和Relay代理检测这两种DAD（双主机检测）技术等。但有一个非常明显的区别就是，CSS目前仅支持两台交换机的集群，而前面介绍的iStack最多可以支持8～9台交换机的堆叠。

5.3.1 CSS基本概念

像本章前面介绍的iStack堆叠技术一样，在建立集群系统之前，每台交换机都是单独的实体，有自己独立的IP地址和MAC地址，对外体现为多台交换机，用户需要独立地管理所有的交换机；集群建立后集群成员对外体现为一个统一的逻辑实体，用户使用一个IP地址对集群中的所有交换机进行管理和维护，如图5-13所示。

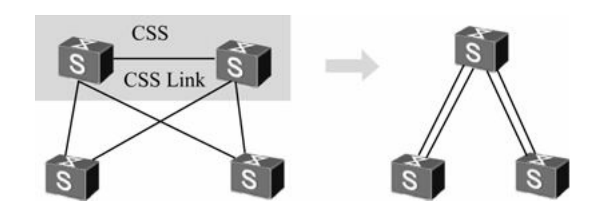


图5-13 CSS集群示意图

在华为CSS集群中主要涉及以下基本概念。

1．角色

目前华为S系列交换机仅支持两台交换机的集群，集群中两台交换机都称为成员交换机，按照功能不同，它们分为以下两种角色（因为CSS中只有两台成员交换机，所以没有iStack堆叠中的“从交换机”角色了）。

（1）主交换机：显示为Master，负责管理整个集群系统。集群系统中只有一台主交换机 。

（2）备交换机：显示为Standby，是主交换机的备份交换机。当主交换机故障时，备交换机会接替原主交换机的所有业务。集群系统中只有一台备交换机 。

2．集群ID

在iStack堆叠中每个成员交换机有堆叠ID，CSS集群中各成员交换机也有对应的集群ID，用来标识和管

理各成员交换机。集群中所有成员交换机的集群ID都是唯一的。

3．集群优先级

因为CSS集群与iStack堆叠同样涉及主交换机的选举，所以各成员交换机也有对应的“集群优先级”属性，用于角色选举过程中确定成员交换机的角色。与iStack优先级一样，CSS集群优先级也是优先级值越大，优先级越高，优先级越高当选为主交换机的可能性越大。因为CSS只有两台成员交换机，所以主交换机的选举就很简单了。

4．集群物理成员端口

集群物理成员端口是指交换机LPU（Line Processing Unit，线路处理单元）单板上专用于集群连接的物理端口。集群物理成员端口用于转发需要跨成员交换机的业务报文或成员交换机之间的集群协议报文。

5．集群端口

集群端口是指用于业务口集群连接方式的逻辑端口，需要和上面介绍的集群物理成员端口绑定。集群的每台成员交换机支持两个集群端口，为CSS-Portn/1和CSS-Portn/2，其中n为成员交换机的集群ID。

5.3.2 CSS特性的产品支持

目前在华为Sx700大系列中，仅S7700和S9700系列支持CSS集群，目前主流应用的S9300、S9300E系列也支持CSS集群，且都只支持两台机的集群 。在S7700、S9300、S9300E和S9700系列交换机中支持集群特性的型号如下。

（1）S7706、S7712（S7706和S7712之间可以混合集群）。

（2）S9306、S9312（S9306和S9312之间可以混合集群）。

（3）S9306E、S9312E（S9306E和S9312E之间可以混合集群）。

（4）S9706、S9712（S9706和S9712之间可以混合集群）。

华为S系列交换机所支持的集群特性包括集群建立、集群的管理和维护、集群快速升级和集群双主检测等。下面具体介绍。

1．集群线缆的连接

在S7700、S9300和S9700系列交换机的集群成员交换机之间的连接方式有集群卡连接和业务口连接两种方式。

（1）集群卡连接

在集群卡连接方式中，每台交换机上必须配置两块同类型的RPU（Route Processing Unit，路由处理单元）主控板，即都是SRUA或SRUB；两台交换机之间可配不同类型的 SRU 主控板，然后在每块 SRU 主控板上插入专门的集群卡。**S7700** 和 **S9300E**系列交换机支持集群卡连接方式；而**S9700**和**S9300E**系列交换机不支持集群卡连接方式 。

在这种集群连接方式中，集群成员交换机之间通过 SRU 主控板上插入的集群卡连接（每块集群卡上有4个集群口）。在两台交换机都有两块主控板的情况下，通过专用的集群电缆QSFP+高速线缆或QSFP+光模块和光纤将这8组集群口按照图5-14规则连接起来。集群口连接规则是固定的，所有集群接口都要插上集群线缆，且不能随意连接 。

（2）业务口连接

在业务口连接方式中，集群成员交换机之间通过LPU单板上的普通业务口连接，无需在SRU主控板上插入专门的集群卡。仅**S7700**、**S9300**、**S9300E**和**S9700**系列支持此种集群连接方式 。

业务口集群连接方式是将 LPU 上的业务口配置为集群物理成员端口后加入逻辑集群端口，然后通过SFP+光模块（如图5-15所示）和光纤或SFP+集群线缆将集群物理成员端口按照图5-16规则连接起来。S7700

和S9700都支持业务口连接方式。对于业务口连接方式，每台交换机可以插上一块或者两块SRU主控板。支持配置业务口连接方式的主控板有SRUA、SRUB和SRUD。

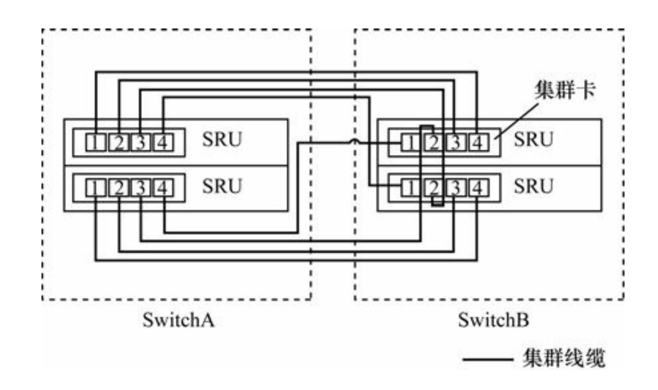


图5-14 集群卡连接规则

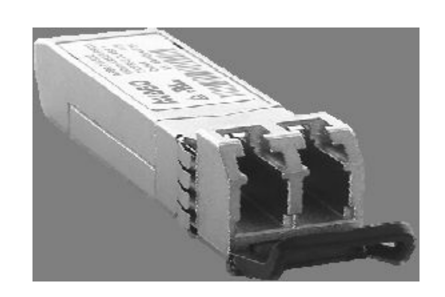


图5-15 SPF+光模块

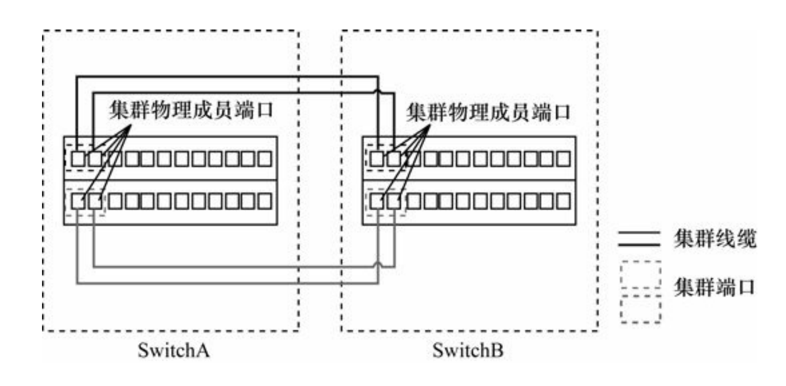


图5-16 业务口连接规则

业务口集群方式具有灵活的组网形式，每块单板最多可配置 32 个集群物理成员端口，提高了集群链路的带宽和可靠性。业务口集群按照链路的分布，又有两种组网形式。

（1）1+0 组网：配置一个逻辑集群端口，物理集群端口分布在一块单板上，依靠一块单板上的集群链路实现集群连接。

（2）1+1 组网：配置两个逻辑集群端口，物理集群端口分布在两块单板上，不同单板上的集群链路形成备份。图5-16就是这样一种组网方式。

为保证集群系统稳定，集群连线时需注意以下几点。

（1）每个逻辑集群端口下加入的物理集群端口数量不限，但是一个逻辑集群端口下的物理集群端口只能与对端交换机的一个逻辑集群端口下物理集群口相连，不允许混连，以避免出现如图5-17所示的连接

方式（SwitchA中的同一个逻辑集群端口下的物理端口与对端SwitchB的两个SRU单板上的物理端口连接了）。

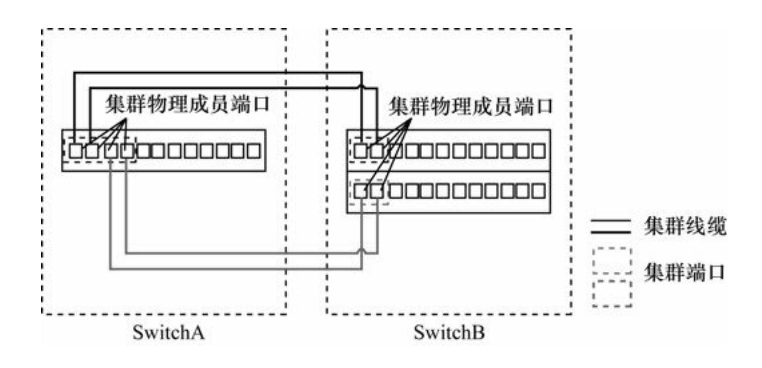


图5-17 错误的连接示意图

（2）在1+1组网中，建议两块单板上的集群链路数量保持一致，并且使用相同端口速率的单板来配置物理集群端口。同时对于S9712或S7712两块单板建议对称分布在主控板的两侧，例如6和7槽位，5和8槽位、……、1和12槽位，而对于S9706或S7706没有这个限制。

2．支持配置业务口集群的单板类型

S9700系列交换机支持业务口集群的单板类型包括EH1D2X40SFC0、EH1D2X40SFC1、EH1D2X16SFC0、EH1D2X16SFC1、EH1D2X08SED4、EH1D2X08SED5、EH1D2L02QFC0、EH1D2L08QFC0和EH1D2X12SSA0。

S9300系列交换机支持业务口集群的单板类型包括LE0DX12XSA00、LE0DX40SFC00、LE0DX16SFC00、LE1D2L02QFC0、LE2D2X08SED4和LE2D2X08SED5。

S9300E系列交换机支持业务口集群的单板类型包括LH2D2X12SSA0、LE0DX40SFC00、LE0DX16SFC00和LH2D2L02QFC0。

S7700系列交换机支持业务口集群的单板类型包括ES0D0X12SA00、ES1D2X16SFC0、ES1D2X40SFC0、ES1D2X08SED4、ES1D2X08SED5和ES1D2L02QFC0。

3．集群建立

在建立CSS集群时，成员交换机间相互发送集群竞争报文，选举出主交换机，负责集群系统的管理。主交换机选举规则如下（依次进行比较，直到选举成功）。

（1）运行状态比较，已经运行的交换机优先处于启动状态的交换机竞争为主交换机。

（2）如果两台交换机都处于启动状态，则进行集群优先级比较，集群优先级高的交换机优先竞争为主交换机。

（3）如果集群优先级也一样，则进行MAC地址比较，MAC地址小的成员交换机优先竞争为主交换机。

（4）如果MAC地址也一样，则进行集群ID比较，集群ID小的成员交换机优先级竞争为主交换机。

选举成功后，如果主、备交换机的软件版本号不一致，则备交换机将同步主交换机的软件版本，复位重启后加入集群系统。集群建立后，集群成员对外体现为一个统一的逻辑实体，用户使用一个IP地址对集群中的所有交换机进行管理和维护。集群系统的IP地址和MAC地址为集群系统首次建立时主交换机的IP地址和MAC地址。

4．集群的管理和维护

CSS集群建立后，所有的成员交换机组成一台虚拟交换机存在于网络中，所有成员交换机的资源由主交

换机统一管理。用户可以通过LPU接口板上的业务口、系统主用主控板上的串口或管理网口登录集群系统，对整个集群系统进行管理和维护。

与本章前面介绍的iStack堆叠一样，在CSS集群建立后，各成员交换机上的接口编号要进行对应的修改，需加上成员 ID 进行区别。对于单台没有运行集群的交换机接口编号采用的格式为槽位号/子卡号/端口号，共三部分；交换机加入集群后接口编号采用的格式为集群ID/槽位号/子卡号/端口号，共四部分 。如交换机没有运行集群时某个接口的编号为GigabitEthernet1/0/1，则当该交换机加入集群后，如果集群ID为2，则该接口的编号将变为GigabitEthernet2/1/0/1。

在集群环境下，业务流量转发与单机环境下不同，跨交换机的转发需要经过交换网 两次，不是直接从集群内部的一台交换机转发到另一台交换机上。对于报文内容的处理没有区别，都需要进行一次上、下行处理。

5．配置文件的备份与恢复

交换机从非集群状态进入集群状态后，会自动将原有的非集群状态下的配置文件进行备份（自动将原有的配置文件加上.bak 的扩展名），以便在去使能集群功能后恢复原有配置。如原配置文件扩展名为.cfg，则备份配置文件扩展名为.cfg.bak，如原配置文件扩展名为.zip，则备份配置文件扩展名为.zip.bak。

在去使能交换机的集群功能时，如果用户想要恢复交换机的原有配置，则可以通过更改备份配置文件名并指定其为下一次启动配置文件，重启交换机即可恢复原有配置。

6．集群分裂

与前面介绍的iStack堆叠一样，集群系统也可能出现分裂现象。在CSS集群建立后，主、备交换机之间定时发送心跳报文来维护集群系统的状态，集群电缆发生故障可能会导致两台交换机之间失去通信，两台交换机之间的心跳报文超时，此时集群系统将分裂为两台独立的交换机，如图5-18所示。

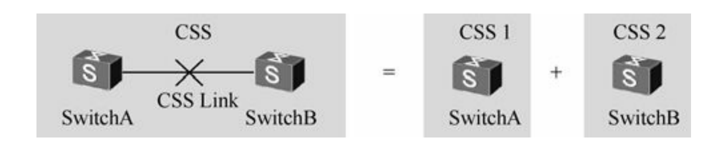


图5-18 CSS分裂示意图

CSS集群系统分裂后，若两台交换机都在正常运行，其全局配置完全相同，会以相同的IP和MAC地址与网络中的其他交换机交互，就会导致IP地址和MAC地址冲突，引起整个网络故障，此时即需要依靠下面将要介绍的集群双主检测（DAD）解决。

7．双主检测

双主检测（Mad Detect，DAD），是一种检测和处理集群分裂的协议，可以实现集群分裂的检测、冲突处理和故障恢复，降低集群分裂对业务的影响。

与 iStack 双主检测一样，CSS 集群的双主检测方式也有“直连检测”和“Relay 代理检测”两种方式。具体参见5.1.2节第9点介绍。

说明

因为集群中的双主检测与本章前面介绍的iStack堆叠中的双主检测配置方法完全一样，只不过这里需要在集群机上配置，而不是在堆叠交换机上配置，故在本节后面不作具体介绍，具体配置步骤参见本章5.2.5节即可。

8．快速升级

集群快速升级提供一种在集群系统的成员交换机软件版本升级过程中不中断当前转发业务的机制，减

少了升级交换机对业务的影响。

在集群进行快速升级时，备交换机将先以新版本重新启动，完成升级，此时数据流量由主交换机转发。在备交换机升级的过程中主交换机将触发集群分裂并变成一个单机集群的系统。在备交换机完成升级后，备交换机升级为主交换机，转发数据流量，此时原主交换机以新版本重新启动，完成升级后成为集群系统的备交换机。

为了确保在集群升级过程中不出现数据流中断的现象，需要确保与集群相连的交换 机使用双线冗余连接方式，也就是分别与主、备交换机都有直接的链路连接，一般是通过跨集群交换机的Eth-Trunk链路来实现的，如图5-19中的其他交换机都与集群中的两台交换机通过Eth-Trunk链路直接的连接。

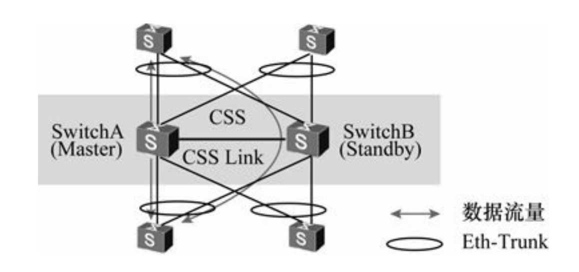


图5-19 CSS集群快速升级时的交换机连接示意图

另外，需要在集群中配置本地流量优先转发功能，使得数据在升级过程中直接从当前担当主交换机角色的交换机上转发。在集群快速升级过程中，会出现集群分裂，会看到集群分裂的告警，但这属正常现象。

5.4 CSS集群配置与管理

在了解了华为S系列交换机中的CSS集群基础知识后，本节要正式介绍CSS集群的配置与管理方法。先了解一下CSS集群配置过程中的一些注意事项和缺省配置。

5.4.1 配置注意事项及缺省配置

1．配置注意事项

在配置CSS集群中要注意以下事项。

（1）建议使用相同端口速率的单板配置业务口集群，不同单板可能会导致跨集群交换机的流量转发不稳定。

（2）对于一块配置物理集群口的单板，在组网时建议不要部署跨集群交换机转发业务，因为来自其业务口的跨集群交换机业务流量本板优先转发，不进行板间负载均衡。

（3）业务口集群支持FSU（Flexible Service Unit，灵活服务单元）子卡，使用FSU子卡时，集群系统中的主控板必须同时插上FSU子卡。

插有FSU子卡，并配有业务口连接方式的集群交换机，在降级到不支持业务口集群的版本时，会出现降级失败，降级前需删除业务口连接方式集群的配置。S9700 从V200R001C01版本开始支持业务口连接方式集群，S7700从V200R002版本开始支持业务口连接方式集群，S9300和S9300E从V200R002版本开始支持业务口连接方式集群。

2．业务口配置为集群物理成员端口后支持的命令

在业务口配置为集群物理成员端口后，该接口仅用来传输集群相关的报文，不能配置业务。同时很多

接口视图下的命令也将屏蔽，但仍然支持以下命令。

● **set flow-stat interval**

● **description** （接口视图）

● **log-threshold**

● **trap-threshold**

● **display interface**

● **display interface brief**

● **display interface description**

● **display counters**

● **reset counters interface**

● **reset counters if-mib interface**

● **set flow-statistics include-interframe**

3．缺省配置

与CSS集群相关的参数缺省配置如表5-5所示。

表5-5 CSS集群相关的参数缺省配置

参数	缺省值
集群使能状态	未使能
集群方式	S9700 和 S9300E 系列交换机仅支持业务口连接模式，缺省均为业务口连接方式，S7700 和 S9300 系列同时支持业务口连接方式和集群卡连接方式，缺省为集群卡连接方式
集群 ID	1
集群优先级	1

5.4.2 CSS集群配置任务

与本章前面介绍的iStack堆叠配置差不多，CSS集群的配置也不是很复杂，最基本的配置总的来说就包括几个方面：使能功能，指定集群端口，配置集群ID和集群优先级，当然还可以有一些其他可选配置任务。针对S7700、S9300、S9300E和S9700系列所支持的集群卡连接方式和业务口连接方式的集群建立流程分别如图5-20的左、右图所示。

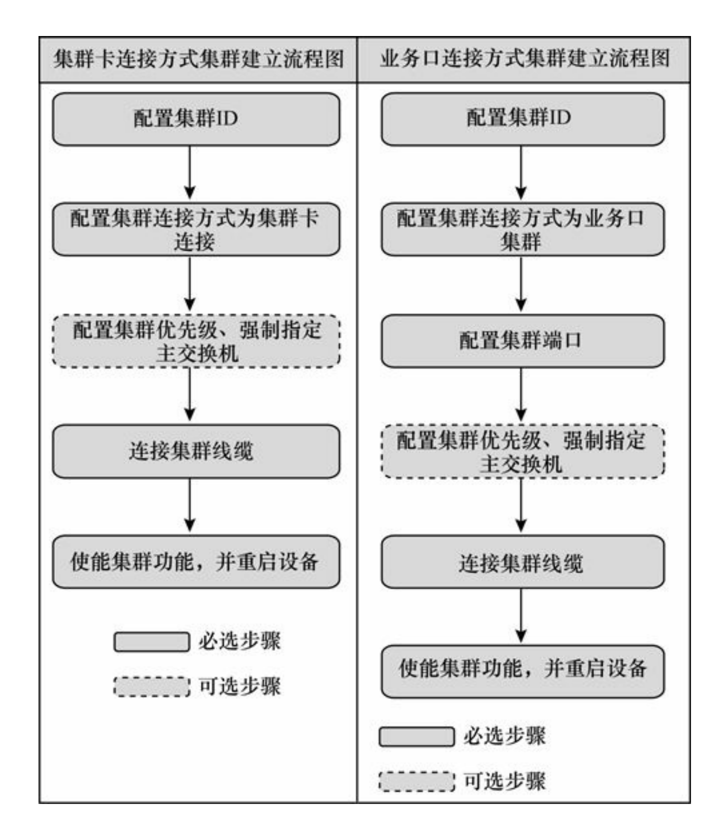


图5-20 两种集群连接方式下的集群建立流程示意图

当重启交换机集群建立后，用户可根据需要选择性配置故障恢复时接口延时 Up 功能、集群系统MAC地址、交换机快速升级等。

下面综合介绍以上两种CSS集群连接方式下的集群基本配置任务，注意这些配置任务中大部分是可选的，而且不同S系列交换机中在配置任务上也有所不同。

1．配置集群ID

集群中的两台交换机拥有不同的集群ID，分别为1或者2，相同ID的两台交换机不能建立集群。缺省情况下，交换机的集群ID都为1，所以在建立集群前，需要手工配置集群中的一台交换机集群 ID 为 2。集群建立后，请勿修改交换机的集群 ID，否则将导致集群分裂。

2．配置集群连接方式

前面介绍了CSS集群成员交换机之间的连接方式有“集群卡连接”和“业务口连接”两种，集群卡连接方式是集群成员交换机之间通过SRU主控板上的集群卡连接；业务口连接方式是集群成员交换机之间通过LPU上的普通业务口连接。两种连接方式互斥，只能以其中一种连接方式组建集群。各系列交换机对这两种连接方式的支持请参见本章5.3.1节相关内容。

3．配置集群端口（仅采用业务口连接方式时需要 ）

在建立业务口连接方式的集群时需要指定单板上的一个或多个端口为集群物理成员端口，并加入逻辑集群端口。但只有配置集群连接方式为业务口连接时才需要执行此项配置任务，集群卡连接方式下不需要配置 。

4．（可选）配置集群优先级

集群优先级主要用于角色选举过程中确定成员交换机的角色，优先级值越大表示优先级越高，优先级越高当选为主交换机的可能性越大。缺省情况下，交换机的集群优先级为1。

5．（可选）强制指定集群主交换机

通常情况下，集群主交换机是在集群系统建立时两台交换机通过竞争产生的，具有不确定性。用户可以通过命令方式强制指定其中某一台交换机作为集群系统的主交换机。但此配置在交换机重启后生效。

6．使能集群功能

关闭两台交换机的电源，连接集群中两成员交换机。当采用集群卡连接方式时按照图5-14所示的连线规则使用专门的集群电缆连接两成员交换机集群卡上的各堆叠端口；当采用业务口连接方式时使用专用集群电缆或者光纤按照图5-16所示的连线规则连接两成员交换机上的各集群物理成员端口。然后开启两交换机的电源，进入系统后分别使能集群功能。然后通过save用户视图命令保存配置，通过rebbot用户视图命令或者手工重启交换机，以使配置生效，如果没有立即重启，则本次配置无效，需重新配置。

7．（可选）配置故障恢复时接口延时Up功能

在交换机集群情况下，如果交换机发生故障会导致集群端口和部分业务口Down。当集群故障恢复后，Down 的端口系统会马上进行配置恢复等流程，此时系统 CPU一般占用率很高。为了避免系统CPU很高时影响正常的业务，可以配置接口延时Up 功能。

8．（可选）配置集群系统MAC地址

集群系统建立后，如果重新启动，或者对主控板拔插更换操作，集群系统的 MAC地址可能会发生变化。用户组建集群后，如果希望集群系统的 MAC 地址保持不变时，可以通过命令将集群系统 MAC 地址设置成某个成员交换机的 MAC 地址，使得集群系统重启后的MAC地址固定为此成员交换机的MAC地址，从而尽量保证集群系统的MAC地址一致。

当配置的 MAC 与当前的系统 MAC 一致时，不需要重启交换机就能生效，否则提示整机重启才能生效。

9．（可选）配置交换机快速升级

用户可以通过命令对集群系统进行版本快速升级，以节约升级交换机的时间，减少因升级交换机对业务带来的影响。此时，与集群相连的交换机使用Eth-Trunk双归连接方式，并且集群Eth-Trunk配置了本地流量优先转发，否则可能会产生数据流量的中断。

5.4.3 配置CSS集群

本节要依据上节介绍的配置任务来介绍各项配置任务下的具体配置方法，具体如表5-6所示。但要区分两种不同的集群连接方式。

表5-6 CSS集群的配置步骤

配置任务	步骤	命令	说明
	1	**system-view** 例如：<HUAWEI> **system-view**	进入系统视图
配置集群ID	2	**set css id** *new-id* [**chassis** *chassis-id*] 例如：[HUAWEI] **set css id 2**	配置交换机的集群 ID。命令中的参数说明如下。 （1）*new-id*：指定生效后的集群 ID，取值为 1 或 2 （2）*chassis-id*：可选参数，指定要修改集群 ID 的成员交换机当前集群 ID，取值为 1 或 2 缺省情况下，交换机的集群 ID 为 1。但如果是在集群状态下配置，且未指定 *chassis-id* 可选参数时，则是对主交换机进行集群 ID 修改
配置集群连接方式	3	**set css mode** { **lpu** | **css-card** } 例如：[HUAWEI] **set css mode css-card**	配置集群成员交换机之间的连接方式。命令中的选项说明如下： （1）**lpu**：二选一选项，指定交换机采用业务口连接方式 （2）**css-card**：二选一选项，指定交换机采用集群卡连接方式 以上两种集群连接方式不兼容，同一时间只能使用一种方式集群，不支持动态切换。在 SRUA/SRUB 主控板环境下，支持两种连接方式，可进行自由选择；但在 SRUD 主控板环境下，只支持业务口集群，交换机不支持该命令 缺省情况下，交换机的集群连接方式为集群卡连接方式

（续表）

配置任务	步骤	命令	说明
配置集群端口（**仅采用业务口连接方式时需要配置**）	4	**interface css-port** *port-id* 例如：[HUAWEI] **interface css-port** 2/1	配置逻辑集群端口，并进入集群端口视图。参数 *port-id* 的格式为：集群 ID/逻辑集群端口号，集群 ID 和逻辑集群端口号只能为 1 或 2。**只有业务口方式连接的集群才支持此命令** 缺省情况下，交换机未配置逻辑集群端口，可用 **undo interface css-port** 命令取消配置的逻辑集群端口。但在取消某个逻辑集群端口前，需要先通过 **shutdown interface** { *interface-type interface-number1* [**to** *interface-type interface-number2*] } &<1-10>命令将集群口下的所有集群物理成员端口关闭。取消某个逻辑集群端口会删除该集群端口下所有集群物理成员端口，如果取消了交换机上的所有的逻辑集群端口，即会造成集群分裂 【注意】一个逻辑集群端口映射到一个单板上，不允许将两个单板上的接口加入同一个逻辑集群口上；一个逻辑集群端口中的物理成员端口只能与另一个逻辑集群端口中的物理成员端口相连，不允许一个逻辑集群端口中的物理成员端口同时与两个逻辑集群端口中的物理成员端口相连
	5	**port interface** { *interface-type interface-number1* [**to** *interface-type interface-number2*] } &<1-10> **enable** [HUAWEI-css-port2/1]**port interface** xgigabitethernet 2/5/0/1 **to** 2/5/0/3 **enable**	配置业务口为集群物理成员端口，并将集群物理成员端口加入以上集群端口中。**只有业务口方式连接的集群才支持此命令，且一个单板上的端口只能加入一个逻辑集群端口** 缺省情况下，接口未配置为集群物理成员端口，可用 **undo port interface enable** 命令还原集群物理成员端口为业务口 【注意】在添加集群物理成员端口时要注意以下事项。 (1) 一个集群物理成员端口只能加入一个逻辑集群端口 (2) 40GE 端口拆分的 XGE 端口不支持加入到集群端口中 (3) 同一个集群端口内的集群物理成员端口所连接的对端集群物理成员端口也必须在同一个集群端口内，不能跨集群端口连接 (4) 还原某集群物理成员端口为业务口时，需首先在逻辑集群端口视图下执行 **shutdown interface** { *interface-type interface-number1*[**to** *interface-type interface-number2*] } &<1-10>命令关闭此集群物理成员端口 (5) 在业务口转换成集群物理成员端口或者集群物理成员端口转换成业务口过程中，端口有可能产生 CRC 错误。建议用户先执行命令 **shutdown** 关闭端口，配置完成后再执行 **undo shutdown** 将端口恢复 Up
（可选）配置集群优先级	6	**quit** 例如：[HUAWEI-css-port2/1] **quit**	退出集群端口视图，返回系统视图

（续表）

配置任务	步骤	命令	说明
（可选）配置集群优先级	7	**set css priority** *priority* [**chassis** *chassis-id*] 例如：[HUAWEI] **set css priority** 100 **chassis** 1	配置交换机的集群优先级。此配置在交换机重启后生效。在两集群机成员同时启动的情况下，优先级高的为集群主交换机；在两成员交换机的启动时间差超过 5s 的情况下，先启动的为集群主交换机。命令中的参数说明如下。 （1）*priority*：指定配置生效后的集群优先级，取值范围为 1～255 的整数。值越大，优先级越高 （2）*chassis-id*：可选参数，指定要修改集群优先级的成员交换机当前集群 ID，取值为 1 或 2 缺省情况下，交换机的集群优先级为 1。但如果是在集群状态下配置，且未指定 *chassis-id* 可选参数时，则是对主交换机进行集群优先级进行修改
（可选）强制指定集群主交换机	8	**css master force** [**chassis** *chassis-id*] 例如：[HUAWEI] **css master force chassis** 1	强制指定成员交换机为集群主交换机。命令中的可选参数 *chassis-id* 用来指定要指定为主交换机的交换机集群 ID，取值只能为 1 或 2。如果不指定此可选参数，则是在当前主交换机上进行操作。此配置在交换机重启后生效 缺省情况下，未强制本机框在集群系统中作为集群主交换机，可用 **undo css master force** [**chassis** *chassis-id*]命令恢复缺省情况
使能集群功能	9	**css enable** 例如：[HUAWEI] **css enable**	使能交换机的集群功能。但在配置此功能前必须按照相应的集群连接方式连接好两交换机。此配置完成后必须立即重启使配置生效，如果没有立即重启，则本次配置无效，需重新配置 缺省情况下，交换机的集群功能处于未使能状态，可用 **undo css enable** [**all** | **chassis** chassis-id]命令去使能该功能。但去使能命令仅在集群使能时可以配置，集群未使能时不可执行；如果选择了二选一可选项 **all**，则表示对两成员交换机的集群功能同时去使能。此时如果有一个交换机操作失败，则两成员交换机都不会去使能
（可选）配置故障恢复时接口延时 Up 功能	10	**css standby port delay** *time* 例如：[HUAWEI] **css standby port delay** 360	配置集群故障恢复接口延时 Up 时间，取值范围是 0～3 600s 在交换机集群情况下，如果交换机发生故障，在集群故障恢复后系统马上进行配置恢复等流程，此时系统 CPU 一般占用率很高。为了避免系统 CPU 占用率很高时影响正常的业务，可以通过本命令配置接口延时 Up 功能 **本命令为覆盖式命令，多次执行后按最后一次配置生效。** 缺省情况下，故障恢复时接口延时 Up 的时间为 0s，即故障恢复时接口不延时 Up，可用 **undo css standby port delay** 命令取消交换机集群时故障恢复接口延时 Up 功能
（可选）配置集群系统 MAC 地址	11	**set css system-mac chassis** *chassis-id* 例如：[HUAWEI] **set css system-mac chassis** 1	设置集群系统的 MAC 地址为指定集群 ID 的成员交换机的 MAC 地址，参数用来指定集群成员交换机的集群 ID，取值只能是 1 或 2 【说明】用户组建集群系统后，如果希望每次集群重新启动后系统 MAC 地址与上一次保持一致时，可以通过本命令将集群系统 MAC 地址设置成集群中某个成员交换机的 MAC 地址，使得集群系统每次重启后的 MAC 地址为此成员交换机的 MAC 地址，从而保证每次重启后的 MAC 地址一致 当配置的 MAC 与当前的系统 MAC 一致时，不需要重启交换机就能生效，否则提示整机重启才能生效

（续表）

配置任务	步骤	命令	说明
（可选）配置交换机快速升级	12	**quit** 例如：[HUAWEI] **quit**	退出系统视图，返回用户视图
	13	**startup system-software** *system-file* **all** 例如：<HUAWEI> **startup system-software** basicsoft.cc **all**	配置所有集群交换机系统下次启动时使用的系统软件文件名。命令中的参数和选项说明如下： （1）*system-file*：指定下次启动时所使用的系统软件文件名，格式为：[*drive-name*] [*path*] [*file-name*]，为4～48个字符，不支持空格，不区分大小写。如果未指定*drive-name*，则此值为缺省的存储器名。系统软件必须以“.cc”作为文件扩展名 （2）**all**：指定所有集群成员交换机都将应用本命令的配置 缺省情况下，没有指定下次启动时使用的系统软件 【注意】所指定的系统软件必须保存至堆叠系统中所有主控板的cfcard根目录下。可将系统软件先上传至主用主控板，再执行**copy** *source-filename destination-filename* **all**命令完成其他所有主控板的复制 缺省情况下，没有指定下次启动时使用的系统软件，可用**undo startup system-software** *system-file* **all**命令取消下次启动时使用的系统软件文件名配置
	14	**css fast upgrade** <HUAWEI> **css fast upgrade**	对集群系统按照上面指定的系统软件进行版本快速升级。用户可通过本命令对集群系统进行版本快速升级，节约升级交换机的时间，减少因升级交换机对业务带来的影响

【示例 1】设置2号成员交换机上5号槽位的接口板上的XGigabitEthernet2/5/0/1接口为集群物理成员端口，并加入逻辑集群端口2/1。

<HUAWEI>system-view

[HUAWEI] interface css-port 2/1

[HUAWEI-css-port2/1] port interfacexgigabitethernet 2/5/0/1 enable

【示例 2】设置下次启动系统软件，进行快速升级。

<HUAWEI> startup system-softwareV200R002C00.cc all

<HUAWEI> css fast upgrade

5.4.4 CSS集群管理

配置好CSS集群后，在使用过程中可使用以下display任意视图命令查看CSS相关信息，监控集群系统运行状态，以保证系统运行正常，并方便在出现故障时准确定位。使用reset用户视图命令清除CSS相关统计信息。

（1）使用display css status [saved] [all | chassis chassis-id] 命令查看系统的集群状态信息，包括成员交换机的集群ID、集群优先级、集群使能状态和集群状态。

（2）使用display css system-mac命令查看集群系统的MAC地址。

（3）使用display css portport-id命令查看集群卡方式集群的指定集群接口的状态统计信息。参数port-id用来指定要查看状态统计信息的集群接口，格式为框号/槽位号/子卡号/端口号，取值根据交换机实际配置选取。

（4）使用display css css-port [saved] [all | chassis chassis-id] 命令可查看业务口集群的逻辑集群端口、集群物理成员端口的配置信息。

（5）使用display css channel [chassis chassis-id | all] 命令可查看集群链路的连线信息以及状态信息。

（6）使用 reset counters css port [port-id] 命令清除集群卡连接方式的集群端口的状态统计信息。可选参数port-id用来指定集群接口编号，格式为框号/槽位号/子卡号/端口号，取值根据交换机实际配置选取。若不指定则表示清除所有集群接口的统计信息。

5.4.5 集群卡连接方式CSS配置示例

在目前主流应用的华为S系列交换机中，S7700 和S9300E系列交换机支持集群卡连接方式的 CSS 集群配置。本示例就以 S7700 系列交换机的集群卡连接方式为例介绍CSS集群的配置方法。在配置这种CSS集群功能之前，需要注意以下事项。

（1）每台成员交换机上必须配置同类型的SRU主控板，即都是SRUA或SRUB，但主控板LE02SRUA VER.A和LE02SRUB VER.A不支持集群。

（2）确认每台成员交换机上每块主控板上都插入了集群卡。

（3）两台成员交换机之间已经用专用的集群电缆连接。

（4）两台成员交换机都能够正常启动。

（5）两台成员交换机的 VRP 系统软件版本一致，当主交换机检测到备交换机版本与其不一致时，会将备交换机版本升级或回退。

本示例拓扑结构如图5-21所示，SwitchA和SwitchB组成集群系统。SwitchC连接用户，并通过Eth-Trunk1连接到集群系统。集群系统通过Eth-Trunk2接入OSPF网络。通过交换机的集群功能增大单台交换机的转发容量，并可提供成员交换机间的备份，提高单台交换机的可靠性。

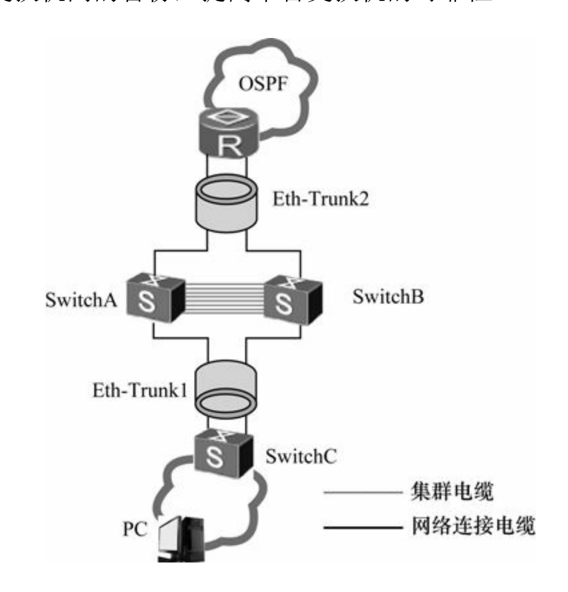


图5-21 集群卡连接方式的CSS集群配置示例拓扑结构

1．配置思路分析

本示例没有特别的要求，只需要先按照5.3.2节图5-14所示连线规则通过专门的集群卡和集群电缆将两台S7700系列成员交换机的堆叠端口连接起来，然后配置CSS集群基本功能，建立CSS集群。根据5.4.2节图5-22左图所示的集群卡连接方式CSS集群建立流程可以得出本示例的基本配置思路如下。

（1）配置两成员交换机的集群ID和集群优先级。

（2）使能两成员交换机的集群功能。

2．具体配置步骤

先设定SwitchA为集群主交换机，这样就只需修改SwitchA的集群优先级为大于1， SwitchB的集群优先级保持缺省优先级1；修改SwitchB的集群ID为2，SwitchA的集群ID保持缺省集群ID为1，然后在两成员交换机上启用集群功能即可。因为S7700、S9300E系列交换机缺省采用集群卡连接方式，所以本示例中无需配置**CSS**集群连接方式 。下面 是具体的配置步骤。

（1）配置SwitchA的集群优先级为200，并使能CSS集群功能。在启用集群功能时会有确认提示，键入

Y或者y确认后即启用了集群功能。配置完成后要立即保存配置，然后重启交换机使配置生效。

```
<HUAWEI>system-view
[HUAWEI] sysname SwitchA
[SwitchA] set css priority 200
Info: CSS config has been changed, need reboot to take effect.
[SwitchA] css enable
Reboot needed to change CSS config. Are you sure this operation and reboot now?[Y/N]y
[SwitchA] quit
< SwitchA >save
< SwitchA >reboot
```

（2）配置SwitchB的集群ID为2，并使能CSS集群功能。配置完成后要立即保存配置，然后重启交换机使配置生效。

```
<HUAWEI>system-view
[HUAWEI] sysname SwitchB
[SwitchB] set css id 2
Info: CSS config has been changed, need reboot to take effect.
[SwitchB] css enable
Reboot needed to change CSS config. Are you sure this operation and reboot now?[Y/N]y
[SwitchB] quit
< SwitchB >save
< SwitchB >reboot
```

重启后在主交换机 SwitchA上通过 display css status命令查看两成员交换机上的CSS状态，验证配置结果。

```
<SwitchA>display css status chassis 1
Property Item          Property Value
Frame ID                1
Priority                255
Enable switch           On
CSS master force        Off
CSS status              master
<SwitchA>display css status chassis 2
Property Item          Property Value
Frame ID                2
Priority                1
Enable switch           On
CSS master force        Off
CSS status              backup
```

5.4.6 业务口连接方式CSS集群配置示例

在目前主流应用的华为S系列交换机中，S7700、S9300和S9700系列交换机都支持业务口连接方式的

CSS集群配置。本示例拓扑结构如图5-22所示，SwitchA和SwitchB两台交换机组成集群系统，两台交换机上的普通业务口XGE1/0/1和XGE1/0/2都加入集群端口。CSS集群通过Eth-Trunk链路与相连的网络交换机实现跨交换机的聚合链路连接。

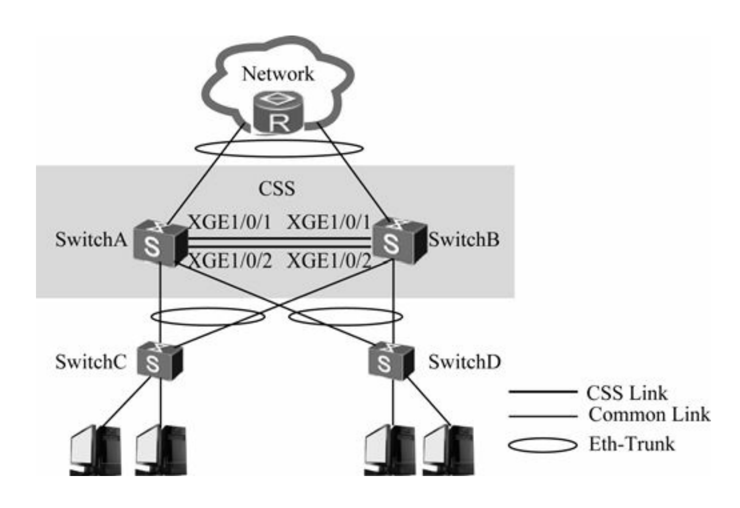


图5-22 业务口连接方式CSS集群配置示例拓扑结构

1．配置思路分析

本示例只是最基本的CSS集群建立配置，没有特别的要求。先按照5.3.2节图5-16所示连线规则通过专门的集群电缆或者SPF+光纤将两台S700系列成员交换机的集群物理成员端口连接起来，然后配置CSS集群基本功能，建立CSS集群。根据5.4.2节图5-20右图所示的业务口连接方式CSS集群建立流程可以得出本示例的基本配置思路如下。

（1）分别配置两成员交换机的CSS集群连接方式为业务口连接方式。

（2）分别配置两成员交换机的集群ID和集群优先级。

（3）配置集群端口，并在集群端口中加入示例中所指定的两个物理成员端口，以增加集群链路的带宽和可靠性。

（4）使能两成员交换机的CSS功能，保存配置，并重启交换机，以使配置生效。

2．具体配置步骤

先设定SwitchA为集群主交换机，这样一来就只需修改SwitchA的集群优先级为大于1，SwitchB的集群优先级保持缺省优先级1；修改SwitchB的集群ID为2，SwitchA的集群ID保持缺省集群ID为1，然后在两成员交换机上配置业务口集群连接方式，向集群端口中添加示例中指定的两个物理成员端口，并启用集群功能即可。

（1）配置SwitchA的集群优先级为200，集群连接方式为业务口连接方式。

```
<HUAWEI>system-view
[HUAWEI] sysname SwitchA
[SwitchA] set css priority 200
[SwitchA] set css mode lpu
```

（2）配置SwitchB的集群ID为2，集群连接方式为业务口连接方式。

```
<HUAWEI>system-view
[HUAWEI] sysname SwitchB
[SwitchB] set css id 2
```

Warning: Modifying the CSS chassis ID will cause interface configuration loss, a

re you sure this operation? [Y/N]:y

[SwitchB] set css mode lpu

（3）在两成员交换机上配置逻辑集群端口（两交换机的集群端口号分别为1/1和2/1），并在集群端口中各自添加XGE1/0/1～XGE1/0/2这两个成员物理端口。

[SwitchA] interface css-port 1/1

[SwitchA-css-port1/1] port interfacexgigabitethernet 1/0/1 to xgigabitethernet 1/0/2 enable

[SwitchB] interface css-port 2/1

[SwitchB-css-port2/1] port interfacexgigabitethernet 1/0/1 to xgigabitethernet 1/0/2 enable

（4）在两台交换机上分别使能集群功能，并分别使用save用户视图命令保存配置，使用rebbot用户视图命令重启两交换机，使配置生效。

[SwitchA] css enable

Warning: The CSS configuration takes effect only after the system is rebooted. T

he next CSS mode is lpu. Reboot now? [Y/N]:y

[SwitchA] quit

< SwitchA >save

< SwitchA >rebbot

[SwitchB] css enable

Warning: The CSS configuration takes effect only after the system is rebooted. T

he next CSS mode is lpu. Reboot now? [Y/N]:y

[SwitchB] quit

< SwitchB >save

< SwitchB >rebbot

（5）两交换机重启后，在主交换机SwitchA上通过display css status命令查看两成员交换机上的CSS状态，通过display css channel命令查看集群端口连线信息，验证配置结果。

<SwitchA>display css status all

Property Item Property Value

Chassis ID 1

Priority 200

Enable switch On

CSS master force Off

CSS status master

CSS mode lpu

Property Item Property Value

Chassis ID 2

Priority 100

Enable switch On

CSS master force Off

CSS status backup

CSS mode lpu

<SwitchA>display css channel

Chassis 1 | Chassis 2

==

Num [Css-port] [Lpu Port] | [Lpu Port] [Css-port]

1 1/1 XGigabitEthernet1/1/0/1 XGigabitEthernet2/1/0/1 2/1

2 1/1 XGigabitEthernet1/1/0/2 XGigabitEthernet2/1/0/2 2/1

5.4.7 CSS集群直连方式DAD配置示例

当集群链路发生故障导致集群分裂时，网络中存在两个配置冲突的集群系统，需要启用双主检测功能，减少集群分裂给网络带来的影响。本示例介绍的是直连方式双主检测的配置方法。

本示例拓扑结构如图5-23所示，SwitchA和SwitchB组成集群系统，SwitchA的集群 ID 为 1，SwitchB 的集群 ID 为 2。配置集群系统的接口 GigabitEthernet1/1/0/5 和GigabitEthernet2/1/0/5直连检测方式的DAD功能。

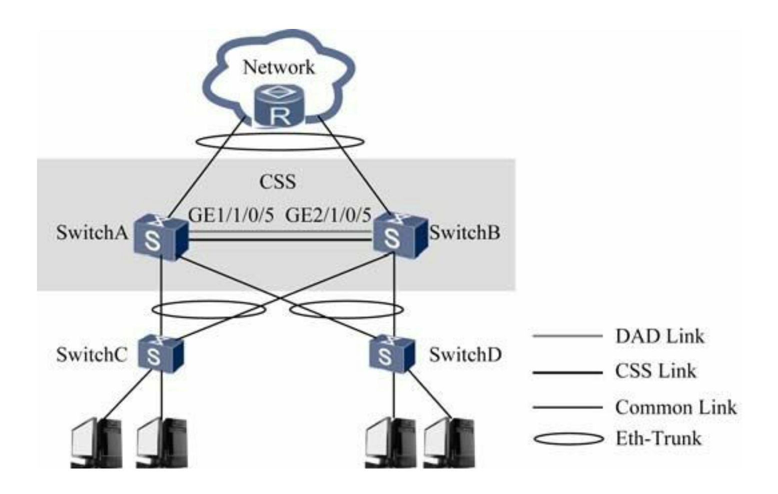


图5-23 直连方式双主检测配置示例拓扑结构

直连方式双主检测的配置方法很简单，只需在两台集群成员交换机中直接连接用于双主检测的端口（本示例分别为GigabitEthernet1/1/0/5和GigabitEthernet2/1/0/5端口）上启用直连方式双主检测功能即可。具体配置方法如下。

（1）在SwitchA上配置接口GigabitEthernet1/1/0/5采用直连检测方式的DAD功能。

<HUAWEI>system-view

[HUAWEI] interfacegigabitethernet 1/1/0/5

[HUAWEI-GigabitEthernet1/1/0/5] mad detect mode direct

Warning: This command will block the port, and no other configuration running on
 this port is recommended. Continue?[Y/N]:y

（2）在 SwitchB 上配置接口 GigabitEthernet2/1/0/5 采用直连检测方式的 DAD功能。

<HUAWEI>system-view

[HUAWEI] interfacegigabitethernet 2/1/0/5

[HUAWEI-GigabitEthernet2/1/0/5] mad detect mode direct

Warning: This command will block the port, and no other configuration running on

this port is recommended. Continue?[Y/N]:y

可通过display mad verbose命令查看集群系统DAD详细配置信息。

<HUAWEI>display mad verbose

Current DAD status: Detect

Mad direct detect interfaces configured:

GigabitEthernet1/1/0/5

GigabitEthernet2/1/0/5

Mad relay detect interfaces configured:

Excluded ports(configurable):

Excluded ports(can not be configured):

5.4.8 CSS集群Relay代理方式DAD配置示例

本示例拓扑结构如图 5-24 所示，SwitchA 和 SwitchB 组成集群系统，SwitchA 和SwitchB通过Eth-Trunk接口与上、下游交换机相连。配置SwitchC作为DAD代理交换机，Eth-Trunk1为集群与SwitchC之间连接使用的接口，采用Relay代理方式实现双主检测。

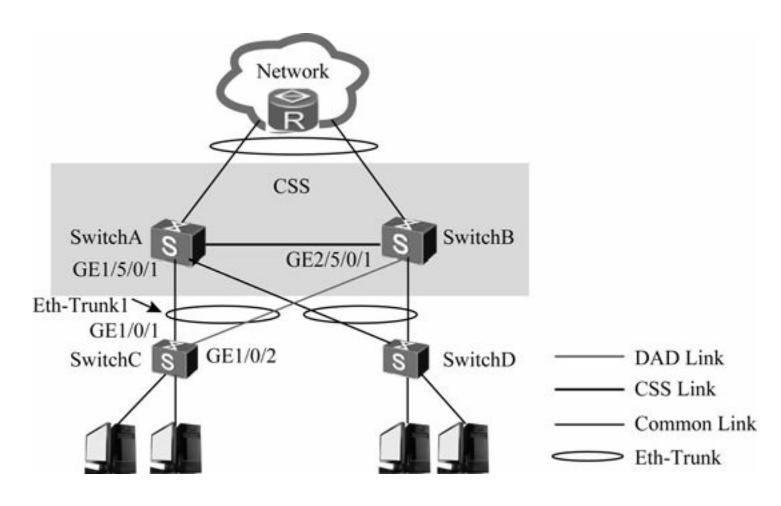


图5-24 CSS集群Relay代理方式双主检测配置示例拓扑结构

采用Relay代理方式双主检测的配置方法是先配置集群系统与代理交换机SwitchC相连的Eth-Trunk接口上分别启用Relay代理检测方式DAD功能。下面是具体的配置步骤。

（1）在集群交换机上配置Eth-Trunk1接口，并使能Relay代理方式双主检测功能（需要在集群主交换机上配置）。

<HUAWEI>system-view

[HUAWEI] interface eth-trunk1

[HUAWEI-Eth-Trunk1] trunkport gigabitethernet 1/5/0/1

[HUAWEI-Eth-Trunk1] trunkport gigabitethernet 2/5/0/1

[HUAWEI-Eth-Trunk1] mad detect mode relay

（2）在代理交换机SwitchC上配置Eth-Trunk1接口，并使能Relay代理方式双主检测功能。

<HUAWEI>system-view

[HUAWEI] sysname SwitchC

```
[SwitchC] interface eth-trunk 1
[SwitchC-Eth-Trunk1] trunkport gigabitethernet 1/0/1
[SwitchC-Eth-Trunk1] trunkport gigabitethernet 1/0/2
[SwitchC-Eth-Trunk1] mad relay
```

配置完成后在任意视图下可通过**display mad verbose** 命令在集群主机上查看集群双主检测详细配置信息，通过**display mad proxy** 命令在代理交换机SwitchC上查看代理信息，验证配置结果。

```
<HUAWEI>display mad verbose
Current DAD status: Detect
Mad direct detect interfaces configured:
Mad relay detect interfaces configured:
  Eth-Trunk 1
Excluded ports(configurable):
Excluded ports(can not be configured):
<SwitchC>display mad proxy
Mad relay interfaces configured:
  Eth-Trunk1
```

第6章　基本VLAN特性配置与管理

说到交换机技术，相信大家都会立即联想起交换机中最重要，也是最常用的一项技术——VLAN（Virtual Local Area Network，虚拟局域网）。从其名称就可以看出，它也是一项局域网（LAN）技术，但它是一项构建虚拟局域网的技术，与物理交换机构建的物理局域网相对。它也是在物理局域网基础之上构建的。

VLAN说起来简单，可真正要掌握并灵活应用它并不那么容易，因为它不仅包括了像基于端口、基于MAC地址、基于子网、基于协议和基于策略等几种VLAN划分方式，VLAN自动注册（GVRP）、VLANIF接口、管理VLAN、VLAN间路由通信等基本特性，又涉及许多VLAN扩展特性及应用，如VLAN聚合（Super-VLAN）、VLAN内用户隔离（MUX VLAN）、VLAN映射（VLAN Mapping）、VLAN多标签封装（QinQ）等。本章先对以上这些基本VLAN特性的配置与管理进行介绍，下章将具体介绍VLAN的一些扩展特性配置与管理。

6.1 VLAN基础

VLAN（虚拟局域网）是可以将一个物理LAN逻辑上划分成多个虚拟LAN的以太网技术。VLAN 划分多个虚拟 LAN 的目的就是要缩小广播域（一个“广播域”就是一个 LAN 网段，即广播数据帧可以到达的节点范围），减小广播数据帧对 LAN 内用户通信的影响。因为一个广播数据帧会在整个LAN内各个节点泛洪发送的，其流量非常大，所占用的带宽资源也非常多。但广播数据帧只能在一个 LAN 中广播，属于二层通信，不能通过路由设备跨网段传输（但可以通过一些代理设备实现跨网段转发，如 ARP 代理），所以只需要把一个大的物理 LAN 划分成多个小的虚拟 LAN 就可以实现缩小广播域的目的，这就是VLAN技术产生的背景。

说明

在LAN通信中有许多通信都会产生广播数据帧的，如ARP在MAC地址寻址时 会产生广播数据帧、DHCP服务器在为客户端自动分配IP地址时也会产生许多广播数据帧。还有许多病毒也会产生许多广播数据帧。

6.1.1 VLAN概述

最终形成的VLAN技术标准 IEEE 802.1Q是于1999年6月由 IEEE委员会正式颁布实施的。随着 20 多年

来的发展，VLAN 技术得到广泛的支持，在大大小小的企业网络中广泛应用，成为当前最主要的一种以太局域网技术。

VLAN 主要用来解决如何将大型网络划分为多个小网络，隔离原本在同一个物理LAN中的不同主机间的二层通信（在物理**LAN**中，各主机是可以直接通过网络体系结构中的第二层，即数据链路层进行通信的，但划分**VLAN**后，不同**VLAN**中的主机是不可以直接通过第二层进行通信的，必须通过第三层，即网络层），以使广播流量不会占据更多带宽资源（因为广播数据帧每复制传播一次都需要消耗一定的带宽和系统资源），同时也提高网段间的安全性，因为广播域缩小了，广播风暴产生的可能性也大大降低了。因为传统共享介质的以太网和交换式以太网中，所有的用户在同一个广播域中（即在同一个LAN中）。

虽然在交换式网络中相对以集线器为集中设备的共享式网络来说缩小了冲突域（共享同一传输介质的节点范围），同时在非全交换全双工模式的以太网络中通过CSMA/CD（Carrier Sense Multiple Access/Collision Detection，载波侦听多路访问／冲突检测）技术提供了冲突避免解决方案，但依然没有解决缩小广播域的问题。LAN内的广播数据帧仍然可以在整个 LAN 内广播，会引起网络性能的下降，浪费宝贵的带宽资源，且其影响随着广播域的增大而迅速增强。此时唯一有效的途经就是重新划分LAN，把单一结构的大LAN划分成相互逻辑独立的小型虚拟LAN，这就是VLAN技术产生的背景。

注意

但在这里不得不说明的是，VLAN的技术基础还是基于以网桥或交换机为集中设备的交换式网络，在以前以集线器为集中设备的共享式网络中 VLAN 标签是不能识别的，因为在集线器的共享网络中数据帧都是以复制方式广播的。只有在交换式网络中才 可能针对具体的目的地址、VLAN标签进行数据转发。

通过将物理LAN划分为多个虚拟的VLAN网段，不仅可以控制不必要的广播数据帧传输，还可以强化网络管理和网络安全。而且VLAN的划分可以突破用户主机地理位置的限制，即不论用户主机实际上是与网络中哪个物理交换机连接，也不管它们所在网络中的物理位置如何，都可以把它们放进同一个虚拟的用户组——VLAN中。相对于物理 LAN 来说具有更好的划分灵活性，因为网络管理员完全可以根据实际应用或管理需求把位于同一物理 LAN 内的不同用户逻辑地划分成不同的 VLAN，而不管这些用户所处的物理位置，连接的是哪个交换机。

图 6-1 所示为一个对分布在各楼层的交换机划分不同 VLAN 的示例。示例中每个VLAN中的成员都分布在不同楼层，而不像物理划分那样仅在一个楼层或者一个部门。所划分的每个VLAN相当于一个小的独立二层交换网络，也就是一个小的广播域。这样每个VLAN中的广播包就只在本地VLAN中广播，而不会传输到其他的 VLAN中去，其影响范围和程度自然就会大大降低。同时如果没有通过三层设备的话，不同VLAN之间不能直接相互通信，这样加强了企业网络中不同部门之间的安全性。

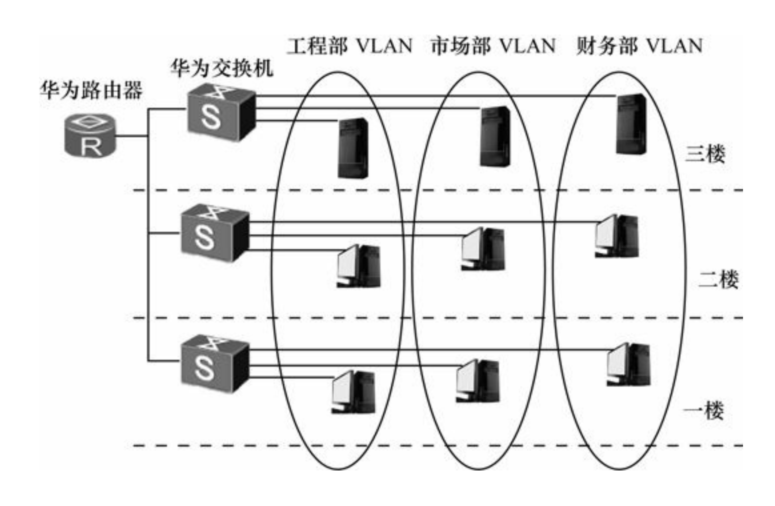


图6-1 VLAN划分示例

6.1.2 理解VLAN的形成原理

前面讲了，网络管理员可按照不同的规则进行VLAN划分，不用考虑各网络用户的实际物理位置。但在VLAN 的配置与使用中，许多读者朋友并没有真正了解 VLAN 的形成原理，导致在出现一些 VLAN 配置和VLAN 路由、桥接故障时无法理解。下面介绍VLAN的形成原理。

1．同一物理交换机中的VLAN形成原理

理解VLAN的形成原理关键就是要理解“虚拟”这两个字。“虚拟”表示VLAN所组成的是一个虚拟，或者说是逻辑 LAN，并不是一个物理 LAN。通过不同的划分规则（具体要依据不同的 VLAN 划分方式而定，将在下节具体介绍）把连接的交换机上的各个用户主机划分到不同的 VLAN 中，同一个交换机中划分的各个 VLAN 可以理解为一个个虚拟交换机。图6-2所示的物理交换机中就划分了5个VLAN，相当于有5个相互只有逻辑连接关系的虚拟交换机。

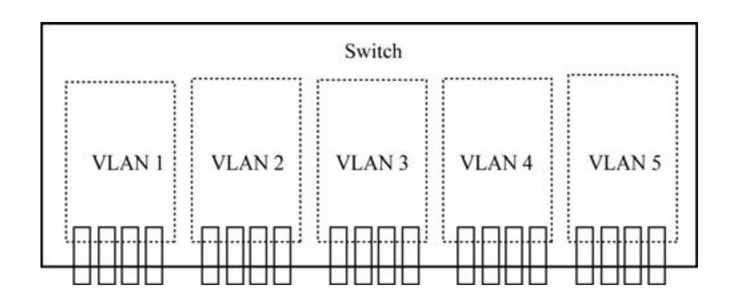


图6-2 一台物理交换机中划分的多个VLAN

其实只要把一个 VLAN 看成一台交换机（只不过它是逻辑意义上的虚拟交换机 ），以前许多问题就比较好理解了，因为虚拟交换机与物理交换机具有许多相同的基本属性。同一物理交换机上的不同VLAN之间就像永远不可能有物理连接，只有逻辑连接 的不同物理交换机一样。既然没有物理连接，那不同VLAN肯定是不能直接相互通信的，即使这些不同VLAN中的成员都处于同一IP网段，因为不同VLAN间的二层通信是隔离了的，只能通过更高的三层进行相互通信。但要注意，这里有一个必须的条件，就是这些不同**VLAN**必须位于同一个物理交换机上 ，如果同处于一个IP网段的不同VLAN位于不同交换机上，则又会有所不同（有时是可以直接互通的），具体将在本节后面介绍。

位于同一VLAN中的端口成员就相当于同一物理交换机上的端口成员一样，不同情况仍可以按照物理交换机来处理。如同一 VLAN 中的各成员计算机可以属于同一个 IP网段，也可以属于不同IP网段，但通常是把属于同一网段的节点划分到同一VLAN中。如果VLAN的各成员计算机都属于同一个IP网段，则没什么，肯定可以相互通信，就像同一物理交换机上连接同一网段的各计算机一样；但如果同一VLAN中的成员计算机是属于不同IP网段，则相当于一台物理交换机上连接处于不同网段的主机用户一样，这时肯定得通过路由或者网关配置来实现相互通信了，即使它们位于同一个VLAN中。

2．不同物理交换机中的VLAN

因为一个VLAN中成员计算机不是依据成员的物理位置来划分的，所以这些成员计算机通常连接在网络中的不同交换机上，这样才更显示出 VLAN 切分的灵活性和实用性。也就是说一个VLAN可以跨越多台物理交换机，这就是VLAN的中继（Trunk）功能。这时就不要总按照物理交换机来看待用户主机的分布了，而要从逻辑的VLAN角度来看待了。图 6-3就不是把它当成两台物理交换机，而是把它当成是五台（VLAN 1、VLAN 2、VLAN 3、VLAN 4和VLAN 5）仅存在逻辑连接关系，但两台物理交换机中相同的 VLAN 间

有相互物理连接关系（就是两台物理交换机间的物理连接）的虚拟交换机了。通常情况下（有特殊情况，具体在下一段介绍），这五个VLAN间的用户是二层隔离的，也就是不能直接互通，仅可以通过网络体系结构中的第三层（网络层）实现互通。

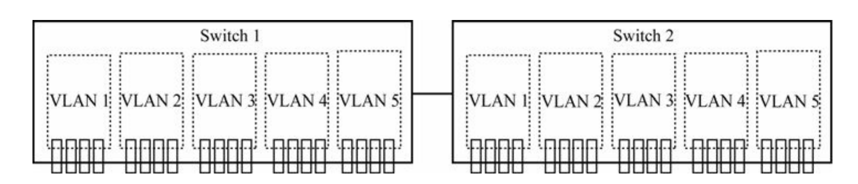


图6-3 不同物理交换机上的相同VLAN

在同一物理交换机上不可能存在两个相同的VLAN，而在不同交换机上可以有相同的VLAN，而且这些不同物理交换机 上的相同**VLAN** 间一般情况下是可以直接互访的，当然这得要求它们都位于同一个IP网段，且在物理交换机连接的端口上允许这些VLAN数据包通过。不仅如此，如果位于不同交换机 上的两个不同**VLAN** 处于同一个**IP**网段，且交换机间连接的两个端口是分别隶属通信双方VLAN的Access端口，或者不带VLAN标签的Hybrid端口，则这两个VLAN间也是可以直接通信的。这就涉及Access、Trunk和Hybrid这三种最基本的二层端口的属性和数据收、发规则了，具体将在本章后面介绍具体二层端口类型时再进行介绍。

【经验之谈】 这里要区分“VLAN 中继”和“中继端口”这两个概念。VLAN 中继是指在一台交换机上的VLAN配置信息可以传播、复制到网络中相连的其他交换机上，在华为交换机上采用的是GVRP（GARP VLAN Registration Protocol，GARP VLAN注册协议）的VLAN自动注册功能来实现的；而Trunk（中继）端口则是指在一个交换机端口允许一个或多个 VLAN 通信到达网络中相连的另一台交换机上相同的 VLAN 中。这是两个不同的概念。有关GVRP的VLAN注册功能及配置方法将在本章6.8节具体介绍。

6.1.3 VLAN标签

前面说了VLAN是二层协议，对应于 IEEE 802.1q标准。这样一来，在二层的数据帧中会打上所属VLAN的标签，这就是 IEEE 802.1q标签（Tag），也就是通常所说的VLAN标签。

1．传统以太网帧格式

传统的以太网数据帧中没有VLAN标签，其帧格式如图6-4所示。各字段说明如下。



图6-4 传统以太网帧格式

（1）前导。前导（Preamble）字段占7个字节，由1和0交互构成（如10101010…），用于使PLS（物理层信号）子层电路与收到的帧达到时钟同步。

（2）帧起始。帧起始（Start-of-Frame Delimiter，SFD）字段占1个字节，前6位也是1和0交互构成，最后两位是连续的1，即10101011，表示一个帧的开始。前导码的作用是使接收端能根据“1”、“0”交互的比特模式迅速实现比特同步，当检测到连续两位“1”（即读到帧起始定界符字段SFD最末两位）时，便将后续的信息递交给MAC子层。

说明

在以上两个字段中，早期的Intel和Xerox公司开发的以太网标准中是把SFD字段并入了Pre字段中，所以那时的MAC帧格式中没有SFD字段。只有后面由IEEE发布的以太网标准中才出现了SFD字段。但Pre和SFD这两个字段只是用来提醒接收端新的一帧到来了，并不计入MAC帧大小中。

（3）目的MAC地址/源MAC地址。目的MAC地址（Destination Address，DA）和源MAC地址（Source Addresses，SA）字段各占 6个字节，分别用于标识接收站点的MAC地址和发送站点的MAC地址。它们可以是单播MAC地址，也可以是组播地址或 广播MAC地址。地址字段最高位为“0”表示单播MAC地址，仅指定网络上某个特定站点；地址字段最高位为“1”、其余位不为全“1”表示组播MAC地址，指定网络上给定的多个站点；地址字段为全“1”，则表示广播MAC地址，指定网络上所有的站点。

（4）长度/类型。“长度/类型”（Length/Type）字段是一个二选一字段，也就是对具体的以太帧来说，它的含义不一样，占两个字节。在Ethernet I和Ethernet II以太网帧中，该字段为“类型”（Type）字段，指出帧中“数据”字段中的数据类型，总大于 1 536（对应的十六进制为 x600）；如果是 IEEE 802.3（包括 Ethernet 802.3 raw、Ethernet 802.3 SAP、802.3/802.2 LLC和 802.3/802.2 SNAP）以太网帧，则该字段为“长度”（Length）字段，值总小于或等于 1 500（对应的十六进制为 x5DC）。在 IEEE 802.3以太网帧中，“数据”字段的长度为 38～1 500个字节。

说明

上面的DA、SA和Length/Type这三个字段组成MAC帧头部。Pre和SFD这两个 字段通常不认为是MAC帧头部的组成部分。

（5）Data。“数据“（Data）字段对于不同的以太网帧所包括的内容也不一样，对于Ethernet I、Ethernet II和Ethernet 802.3 raw以太网帧，它就是从网络层来的数据包；而对于Ethernet 802.3 SAP、802.3/802.2 LLC和 802.3/802.2 SNAP以太网帧，则是LLC帧全部内容，包括LLC帧头和来自网络层的数据包。也正如此，“数据字段”长度范围也各不一样，具体如下。

● Ethernet I、Ethernet II帧“数据字段”长度范围为 46～1 500个字节。

● Ethernet 802.3 raw帧“数据字段”长度范围为 44～1 498个字节。

● Ethernet 802.3 SAP和 802.3/802.2 LLC帧“数据字段”长度范围为 43～1 497个字节。

● Ethernet 802.2 SNAP帧“数据字段”长度范围为 38～1 492个字节。

整体而言，该字段长度范围为 38～1 500个字节。但无论怎样，总的MAC帧长度最小为 64个字节，最长为 1 518个字节（不包括“前导”字段和“帧起始字段”），如果不够64个字节时，要在“数据”字段中加上PAD填充字段。

注意

这里所说的38～1 500个字节长度是在没有经过 IEEE 802.1q VLAN协议重封装时的长度范围，如果封装了VLAN协议，则因为VLAN标签占用了4个字节，所以就整个以太网帧来说，“数据”字段的取值范围就为34～1 500个字节。有关 IEEE 802.1q VLAN协议将在本章后面具体介绍。

（6）Frame Check Sequence。“帧校验序列”（FCS）字段占4个字节，包括32位的循环冗余校验（CRC）值，由发送端对MAC帧自DA字段到Data字段间（不包括Pre和SFD这两个字段）的二进制序列生成校验和，然后通过接收端对所接收的帧的以上部分重新计算，看两次校验的结果是否一样即可以得出所检验的帧在传输过程中是否已被破坏。

2．802.1q帧格式

IEEE 802.1q是虚拟桥接局域网的正式标准，对传统的Ethernet帧格式进行了修改，在“源MAC地址”字段和“长度/类型”字段之间插入了一个4字节的“802.1q Tag”字段。而这个“802.1q Tag”字段又包括了TPID、PRI、CFI和VLAN ID四个子字段，如图6-5所示。

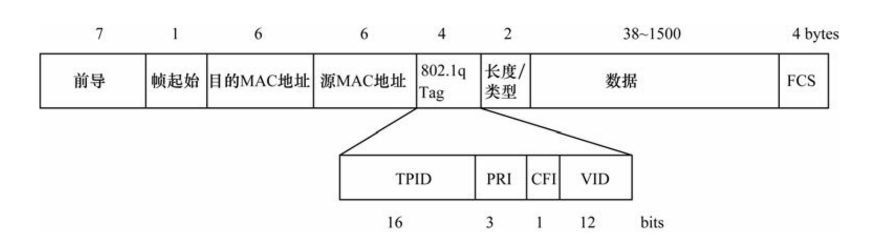


图6-5 802.1q帧格式

"802.1q Tag"字段所包括的 4个子字段的说明如下。

（1）TPID："Tag Protocol Identifier"（标签协议标识符）字段，占两个字节（16位），表明这是一个添加了 IEEE 802.1q标签的帧（区别于未加VLAN标签的帧），值固定为0x8100（表示封装了 IEEE 802.1q VLAN协议）。如果不支持 802.1q的设备（如用户主机、打印机等终端设备就不支持）收到这样的帧，就会将其丢弃。

（2）PRI：Priority（优先级）字段，占3位，表示0～7八个优先级（值越大，优先级越高），主要用于当交换机阻塞时，优先发送哪个数据帧，也就是 QoS（服务质量）的应用，是在802.1p规范中被详细定义的。

（3）CFI："Canonical Format Idicator"（标准格式指示器）字段，占 1位，用来兼容以太网和令牌环网。用来标识 MAC 地址在传输介质中是否以标准格式进行封装，取值为0表示MAC地址以标准格式封装，为1表示以非标准格式封装，缺省取值为0，在以太网中该值总为0，表示以标准格式封装MAC地址。

【经验之谈】 在这里要介绍一个绝大多数图书和文章都没有介绍的问题，那就是什么是"标准格式MAC地址"，什么是"非标准格式MAC地址"。其实，以太网（IEEE 802.3）和令牌总线网（IEEE 802.4）在传输介质中发送MAC地址字节的顺序是从低到高（也就是我们平时写MAC地址格式中的从右到左顺序），而令牌环网（IEEE 802.5）和IEEE 802.6标准中MAC地址字节在传输介绍中的发送顺序则相反，是从高到低的顺序。MAC地址字节的发送顺序也对应MAC地址字节的封装顺序，把 IEEE 802.3和 IEEE 802.4标准中的MAC地址封装顺序称之为"标准格式"（canonical form），而把 IEEE 802.5和 IEEE 802.6标准中的MAC地址封装顺序称之为"非标准格式"（non-canonical form）。例如一个MAC地址为 12-34-56-78-9A-BC，以标准格式发送时，则它的比特次序为 01001000 00101100 01101010 00011110 01011001 00111101，但是以非标准格式发送时，它的比特次序是00010010 00110100 01010110 01111000 10011010 10111100（注意，比较每个字节可以发现，它与前面的标准格式是次序相反的）。

（4）VID："VLAN Identified"（VLAN标识）字段，占 12位，指明VLAN的 ID，取值范围为 0～4 095，共4 096个，但由于 0和4 095为协议保留取值，所以VLAN ID的实际有效取值范围是 1～4 094。每个进入支持 802.1q协议的交换机发送出来的数据包都会包含这个域，以指明自己属于哪一个VLAN。

6.1.4 主要VLAN特性及产品支持

目前华为交换机中所支持的VLAN特性主要包括VLAN划分、VLAN注册、VLAN 间通信、VLAN聚合、MUX VLAN、VLAN映射、Voice VLAN、管理VLAN。当然并不是所有型号交换机都支持以上全部的VLAN 特性，具体将在下面说明。在此仅简单地介绍这些 VLAN 主要特性，因为这些特性的具体配置与应用将在本章或下章全面介绍。

1．VLAN划分

要使用VLAN技术，首先就要创建所需的VLAN，然后把各用户计算机划分到这些不同的VLAN中，这就是VLAN划分特性。

VLAN 最基本的配置是划分 VLAN，VLAN 划分成功后即可实现不同 VLAN 内用户的二层隔离，达到缩小广播域的目的。华为交换机可支持以下 5 种 VLAN 划分方式：基于端口划分、基于MAC地址划分、基于子网划分、基于协议划分、基于策略划分。

在以上这5种VLAN划分方式中又可归总为“静态VLAN划分”和“动态VLAN划分”两大类。所谓“静态VLAN 划分”方式就是指连接用户的交换机端口被固定地划分到一个特定的VLAN中。所支持的VLAN划分方式只有“基于端口划分”方式。而“动态 VLAN 划分”方式中连接用户计算机的端口不是固定地划分到某一特定VLAN中，而是根据用户主机的MAC地址、IP地址、网络层协议或者MAC地址+IP地址或+交换机端口的组合策略来灵活地加入到不同的VLAN中。除了“基于端口划分”方式外的其他4种VLAN划分方式都是属于这种类型。但无论是静态VLAN划分方式，还是动态VLAN划分方式所划分的VLAN均属于静态VLAN。在本章后面 6.8节介绍的，通过GVRP（GARP VLAN注册协议）协议动态注册的VLAN属于动态VLAN。

以上5种VLAN划分方式的划分方法说明、主要优缺点、应用场景及华为S系列交换机产品的支持情况如表6-1所示。

表6-1 各种VLAN划分方式的比较

VLAN划分方式	划分方法	优点	缺点	应用场景	华为交换机的支持
基于端口划分	根据用户所连的二层端口（不能是三层端口）进行划分。网络管理员给交换机的每个二层端口配置不同的 PVID（Port Default VLAN ID，端口缺省VLAN ID），即一个端口缺省属于的 VLAN 当一个数据帧进入交换机端口时，如果没有带 VLAN 标签，则该数据帧就会被打上端口的 PVID；如果进入的帧已经带有 VLAN 标签，那么交换机不会再增加 VLAN 标签，即使端口已经配置了 PVID 对 VLAN 帧的收、发处理由二层端口类型决定，具体将在下节介绍	配置过程简单，是最常用的 VLAN 划分方式	配置不够灵活，当 VLAN 中的成员所连接的端口发生变化时需要重新配置 VLAN。这对于拥有众多移动用户的网络来说，网络管理者将会花费更多的时间进行维护	适用于规模大、安全需求不高的场景中	除 S1700 系列非网管型交换机外，其他所有华为 S 系列交换机都支持

（续表）

<table>
<tr><th>VLAN划分方式</th><th>划分方法</th><th>优点</th><th>缺点</th><th>应用场景</th><th>华为交换机的支持</th></tr>
<tr><td>基于MAC地址划分</td><td>根据用户主机网卡MAC地址进行划分。网络管理员需事先配置MAC地址和VLAN ID映射关系表，如果交换机收到的是Untagged（不带VLAN标签）帧，则依据该映射表在帧中添加对应的VLAN ID</td><td>用户在变换物理位置时，不需要重新划分VLAN。提高了终端用户的安全性和接入的灵活性</td><td>网络管理者需要事先将归属到指定VLAN的终端设备MAC地址配置到交换机上。这对拥有大量终端的网络来说，初始配置时，配置工作量较大</td><td>适用于安全和移动性需求较高的场景中</td><td>除S1700系列非网管型交换机、S2700-SI、S2710-SI系列交换机外的其他所有华为S系列交换机</td></tr>
<tr><td>基于子网划分</td><td>根据用户主机网卡IP地址所在IP网段进行划分。网络管理员需事先配置IP地址和VLAN ID映射关系表，如果交换机收到的是Untagged帧，则依据该映射表在帧中添加对应的VLAN ID</td><td rowspan="2">基于子网划分VLAN和基于协议划分VLAN统称为基于网络层划分VLAN。基于网络层划分VLAN，不但大大减少了人工配置VLAN的工作量，同时保证了用户自由地增加、移动和修改</td><td>交换机需要解析源IP地址并进行相应转换，导致交换机响应速度慢</td><td>适用于对安全需求不高，对移动性和简易管理需求较高的场景中</td><td rowspan="3">除S1700、S2700系列外的其他所有华为S系列交换机</td></tr>
<tr><td>基于协议划分</td><td>根据用户主机运行的网络层协议类型进行划分。网络管理员需要事先配置以太网帧中的“协议”字段和“VLAN ID”字段的映射关系表，如果交换机收到的是Untagged帧，则依据该映射表在帧中添加对应的VLAN ID</td><td>交换机需要分析各种协议的地址格式并进行相应转换，导致交换机响应速度慢</td><td>目前支持AppleTalk、IPv4、IPv6、ipx等网络层协议划分VLAN</td></tr>
<tr><td>基于策略划分VLAN</td><td>根据用户安全策略进行划分。主要包括基于MAC地址+IP地址组合策略和基于MAC地址+IP地址+端口组合策略两种
只有符合条件的终端才能加入指定VLAN。符合策略的终端加入指定VLAN后，严禁修改IP地址或MAC地址，否则会导致终端从指定VLAN中退出</td><td>安全性非常高，可禁止用户改变IP地址或MAC地址。相较于其他VLAN划分方式，基于策略划分VLAN是优先级最高的VLAN划分方式</td><td>针对每一条策略都需要手工配置，在VLAN较多时工作量很大</td><td>适用于规模小，且对安全和移动性需求非常高的场景中</td></tr>
</table>

说明

如果一交换机上同时配置了以上多种方式划分的VLAN，则对于具体的用户主机来说将按以下从高到低的优先级顺序采用最终的VLAN划分方式：基于匹配策略划分VLAN→基于MAC地址划分VLAN和基于子网划分VLAN→基于协议划分VLAN→基于端口划分VLAN。

从以上顺序可以看出，基于MAC地址划分VLAN和基于子网划分VLAN拥有相同的优先级，缺省情况下优先基于MAC地址划分VLAN。但是可以通过命令改变基于MAC地址划分VLAN和基于子网划分VLAN的优先级（S5700LI和S5700S-LI不支持），从而使同时满足这两种划分方式的用户主机决定优先采用的VLAN划分方式。另外，虽然基于端口划分VLAN的优先级最低，但是最常用的VLAN划分方式；基于匹配策略划分VLAN的优先级最高，但是最不常用的VLAN划分方式，因为配置复杂。

有关以上这五种VLAN划分方式的配置方法将在本章6.2～6.6节介绍，有关VLAN划分优先级的配置将在本章6.3节和6.4节具体介绍。

2．VLAN间通信

前面说了，一般情况下不同VLAN是不能直接进行通信的，因为VLAN间缺省是二层隔离的。但有时又需要不同VLAN中的用户进行通信，为此华为S系列交换机提供了3种实现VLAN间通信的解决方案：三层VLANIF接口方案、三层以太网子接口方案和VLAN Switch（VLAN交换）方案。

前两种是通过网络体系结构中的第三层——网络层功能来实现的，通过在各VLAN中的用户主机中配置指向所属VLAN的VLANIF接口或者对应的三层以太网子接口的IP地址作为网关，然后利用三层交换机的

IP路由功能实现交换机内部不同VLAN间的三层互通，或者利用其他路由功能在VLAN间实现三层互通。最后的VLAN Switch方案是通过VLAN标签替换方法来实现的。

本项VLAN特性中的三层VLANIF接口方案，除了S1700系列外的其他所有华为S系列交换机均支持，而三层以太网子接口方案和VLAN Switch方案仅在S7700、S9300、S9300E和S9700等高端系列交换机中支持（**S5700HI**和**S5710EI**系列交换机虽然也支持三层以太网子接口，但不能为子接口配置**IP**地址 ）。有关VLAN间的通信原理和具体的配置方法将在本章6.9节介绍。

3．管理VLAN

管理VLAN是一种特殊的VLAN，是专门用来为用户提供远程设备管理（通过管理VLANIF的IP地址）的VLAN，其中只有管理流（也就是进行各项网络管理的数据流）而无业务流（即用户的网络应用数据流），所以在管理VLAN中不能有业务用户主机。

本项VLAN特性除了S1700系列外的其他所有华为S系列交换机均支持。有关管理VLAN的配置将在本章6.10节介绍。

4．VLAN聚合

VLAN聚合（VLAN aggregation）就是通常所说的Super VLAN技术。它是在一个主VLAN下包括多个同处一个IP网段的从VLAN，但只需要为主VLANIF接口配置IP地址，各从VLAN中的用户主机可以主VLAN的VLANIF接口的IP地址作为缺省网关实现三层互通。很显然它是为了解决传统 VLAN 中要实现不同 VLAN间相互通信必须为每个VLANIF接口配置一个IP地址，造成一定程度上IP地址资源浪费的问题，同时又可实现不同VLAN间的相互通信的目的。

本项VLAN特性除了S1700系列和S2700SI系列之外的其他所有华为S系列交换机均支持。有关VLAN聚合的工作原理和具体的配置方法将在下章介绍。

5．MUX VLAN

MUX VLAN是一种可实现在VLAN内部各成员间二层隔离的技术。传统VLAN仅隔离不同VLAN中用户的二层通信，同一VLAN内的各用户是可以直接进行二层通信的。 但有时又希望在 VLAN 内部的某个成员（如存在不安全因素的用户主机，或者需要特别保护的主机，如某些服务器，或者个别用户的计算机等）与其他成员进行隔离，这时就可以采用MUX VLAN技术来实现，主要是想达到安全保护，或者用户授权的目的。

本项VLAN特性除了S1700、S2700系列之外的其他所有华为S系列交换机均支持。有关MUX VLAN的工作原理和具体的配置方法也将在下章介绍。

6．VLAN映射

传统的VLAN技术可有效地控制广播域范围、隔离不同VLAN中用户间的二层通信。但是由于一些低端交换机不支持全范围（1～4 094）的VLAN ID，而是类似1～512的有限范围，而且还有一些VLAN ID被保留，这样一来就可能导致用户网络中的VLAN ID与公网VLAN ID冲突。于是就产生了一种可以实现在用户VLAN ID和运营商VLAN ID之间相互转换的VLAN技术——VLAN映射（VLAN Mapping）。VLAN映射通过替换数据帧中的VLAN标签来实现用户VLAN与运营商VLAN的相互映射，使用户业务按照运营商的网络规划进行传输。

本项VLAN特性除S1700、S2700SI系列之外其他华为所有S系列交换机都支持，但有些机型不支持全部的 VLAN 映射方式。有关 VLAN 映射的工作原理、具体的配置方法和华为交换机对各映射方式的支持将在下章介绍。

7．Voice VLAN

Voice VLAN（语音VLAN）是相对传统数据VLAN而言的，它是针对不同类型的用户话音流进行划分

的，并能自动修改话音数据帧的优先级，以提高话音数据的传输质量。

本项VLAN特性除了S1700非网管型系列和S2700SI系列外的其他所有华为S系列交换机均支持。因为Voice VLAN相对比较少用，故本书不作具体介绍。

8．VLAN透传

VLAN 透传特性可以使交换机直接透明传输指定 VLAN 内的数据帧，不上送 CPU处理，也就不用去识别数据帧中的VLAN标签，从而提高了转发效率。

本项VLAN特性在S5710EI、S5700HI、S6700、S7700、S9300和S9700系列华为交换机中均支持。具体也将在下章介绍。

9．QinQ

QinQ（802.1Q-in-802.1Q）协议是基于 IEEE 802.1Q技术的一种二层隧道协议。在公网中传递的帧一般有两层802.1Q标签（一个公网VLAN标签，一个私网VLAN标签）。QinQ的核心思想是将用户私网VLAN标签封装在公网VLAN标签中，这样报文带着两层VLAN标签穿越网络运营商的骨干网络，从而为用户提供一种较为简单的二层VPN隧道。

本特性除S1700系列交换机外，其他所有华为S系列交换机均支持，当然不同系列所支持的功能不完全一样，具体也将在下章介绍。

6.2 基于端口划分VLAN

基于端口 VLAN 划分方式是最常用，也是最简单的 VLAN 划分方式，是把交换机端口静态地划分到某一个或多个具体的 VLAN 中，是一种静态 VLAN 划分方式。但要注意的是，因为VLAN是二层协议，所以仅可以把二层以太网端口（包括物理以太网端 口和**Eth-Trunk**聚合链路口 ）划分到VLAN中。虽然华为交换机中的以太网端口可以转换成三层模式，但仍不能直接配置**IP**地址 ，有关以太网端口配置参见本书第4章。下面绍S系列交换机中的二层以太网端口类型。

6.2.1 二层以太网端口

在华为交换机中主要包括 Access（访问）、Trunk（干道）和 Hybrid（混合）、QinQ这4种二层以太网端口。在本节介绍的基于端口VLAN划分方式中可以把前3种二层交换机以太网端口加入特定的VLAN中，但是其他所有**VLAN**划分方式都只能添加**Hybrid**类型端口 。QinQ端口仅用于支持QinQ协议，不能用于VLAN划分。

1．二层以太网端口类型

下面具体介绍华为交换机的Access、Trunk、Hybrid和QinQ这4种二层以太网端口的基本特性和数据帧收、发规则。

（1）Access端口。Access端口主要是用来连接用户主机 的二层以太网端口。它有一种最主要的特性是仅允许一个**VLAN** 的帧通过，反过来也就是Access端口仅可以加入一个VLAN中，且Access端口发送的以太网帧永远是**Untagged** （不带标签）的。

（2）Trunk端口。Trunk端口是用来连接与其他交换机 的二层以太网端口。它的最主要特性是允许多个 **VLAN** 的帧通过，并且所发送的以太网帧都是带标签的，除了发送VLAN ID与PVID（Port Default VLAN ID，端口缺省VLAN ID）一致的VLAN帧。

（3）Hybrid端口。Hybrid端口可以说是以上Access端口和Trunk端口的混合体，它们具有共同的特性，是一种特殊的二层以太网端口。正因如此，Hybrid端口既可以连接用户主机，又可以连接其他交换机、路

由器设备 。同时 Hybrid 端口又允许一个或多个**VLAN** 的帧通过，并可选择以带标签或者不带标签 的方式发送数据帧。

（4）QinQ端口

QinQ端口是专用于**QinQ**协议的二层以太网端口 。它可以给数据帧加上双层VLAN标签，即在原来标签的基础上，给帧加上一个新的标签，从而可以支持多达4 094×4 094个VLAN，满足企业用户网络对VLAN数量更高的需求。S1700和S2700SI不支持QinQ类型端口。

【经验之谈 】 虽然从理论上讲交换机与交换机、交换机与路由器连接之间的链路也可以是**Access**类型的，但在实际的组网应用中通常是带标签类型的 ，可以是Trunk类型，也可以是带标签的 Hybird 类型。一方面是因为不同网络设备间的通信通常是包含多个VLAN间的通信，而Access类型端口仅允许一个VLAN的数据通过，肯定不行；另一方面，Access类型和不带标签的Hybrid类型在发送数据时是不带标签的，这样一来，对端设备接口接收到来自本端设备任何VLAN的数据后都将打上该接口的PVID所对应的VLAN标签，并且被错误地转发到该VLAN中，这显然不符合实际需求，最终造成无法正常通信。

另外，对于连接用户 PC 机、服务器主机或者傻瓜式二层交换机设备的端口仅可以是Access类型或者不带标签的Hybrid类型，因为这些设备不能识别带有VLAN标签的数据帧，而这两种类型端口在发送数据时正好是不带VLAN标签的。傻瓜式二层交换机设备所连接的所有设备都将加入到对端交换机端口所加入的一个VLAN中。

2．二层以太网端口的缺省VLAN

以上Access、Trunk和Hybrid三种类型二层以太网端口都可以配置一个缺省VLAN， 对应的VLAN ID为PVID。但端口类型不同，其缺省VLAN的含义也有所不同。Aceess端口的缺省 VLAN 就是 Access 端口所加入的 VLAN，因为 Aceess 端口只能加入一个VLAN。但Trunk和Hybrid端口的缺省VLAN需要通过命令配置指定，因为它们都相当于加入了多个VLAN，缺省都是VLAN1。有关缺省VLAN的具体配置方法将在本章后面介绍Trunk和Hybrid端口配置时会有所体现。

3．二层以太网端口的数据收发规则

以上Access、Trunk和Hybrid三种类型二层以太网端口在接收和发送数据帧时对帧的处理规则也是不同的，而这些规则直接影响着数据通信的成败，一定要记住。具体如表6-2所示。

【经验之谈 】 这里所说的数据帧“收”是指交换机端口接收从对端设备发来的数据帧 ，而不是接收从交换机内部的另一个端口发来的数据帧，因为在交换机内部传输的数据帧都是带有**VLAN**标签的 ，无论是从哪种交换机端口发来的数据帧。同理，这里所说的数据帧“发”是指从交换机端口向对端设备发送数据帧 ，而不是指本地交换机中一个端口向另一个交换机端口发送数据帧。这一点要特别注意，否则很难理解这些端口的数据接收、发送规则。

表6-2 二层以太网端口数据帧处理规则

端口类型	收到不带 VLAN 标签的帧的处理规则	收到带 VLAN 标签的帧的处理规则	发送帧时的处理规则	用途
Access 端口	接收该帧，并打上该端口所加入 VLAN 的 VLAN 标签	当帧中的 VLAN ID 与端口加入 VLAN 的 VLAN ID 相同时，接收该帧，否则丢弃该帧	当帧中的 VLAN 标签与该端口的 PVID 相同时，则去掉帧中的标签，然后发送该帧，否则丢弃该数据帧。**Access 端口所发送的帧总是不带 VLAN 标签的**	端口只能属于 1 个 VLAN，用于设备与计算机直接连接
Trunk 端口	在帧中打上该端口的缺省的VLAN标签，当此缺省 VLAN ID 在该端口允许通过的 VLAN ID 列表里时，接收该帧，否则丢弃该帧	当帧中的 VLAN ID 在该端口允许通过的 VLAN ID 列表里时，接收该帧，否则丢弃该帧	当帧中的 VLAN ID 与该端口的缺省 VLAN ID 相同，且是该端口允许通过的 VLAN ID 时，则去掉帧中的 VLAN 标签后再发送该帧 当帧中的 VLAN ID 与该端口的缺省 VLAN ID 不同，但仍是该端口允许通过的 VLAN ID 时，保留帧中原有 VLAN 标签并发送该帧 当帧中的 VLAN ID 不是该端口允许通过的 VLAN ID 时，不允许发送	端口允许多个 VLAN 通过，可以接收和发送多个 VLAN 的帧，一般用于网络设备之间的连接
Hybrid 端口			当帧中的 VLAN ID 是该端口允许通过的 VLAN ID 时，则发送该帧（**不管帧中的 VLAN ID 与该端口的缺省 VLAN ID 是否相同**），但可以通过命令配置发送时是否携带原有的 VLAN 标签（**通常只有在与主机连接的链路不需要带 VLAN 标签**） 当帧中的 VLAN ID 不是该端口允许通过的 VLAN ID 时，丢弃该帧	端口允许多个 VLAN 通过，可以接收和发送多个 VLAN 的帧，且既可以用于网络设备之间的连接，也可以用于网络设备与用户设备之间的连接

【经验之谈 】这里有一个大家争论得比较多的一个问题，那就是帧到达Access端口时交换机是否会为帧打上 VLAN 标签。因为许多人认为 Access 端口所发送的帧都是不打 VLAN 标签的，所以有人认为 Access 端口在帧中打上标签是没有任何意义的，也就认为Access端口不会在帧中打上标签。其实这是错误的。

虽然 Access端口向对端设备发送数据时是不带 VLAN 标签的，但是数据到了交换机后还需要一个转发过程，在交换机内部传输中所有数据都是带有VLAN标签的（当然这要求交换机支持 VLAN 才行），而且也有许多时候数据不是直接转发到目的节点的，而是需要在交换机上进行一些处理（如基于VLAN的策略路由、基于VLAN的ACL等），或者进行端口镜像等管理工作，这些都需要识别这些数据是来自哪个VLAN的用户，毕竟在一个交换机的这么多端口上可能分属于不同的VLAN。

6.2.2 二层以太网链路

上节介绍的Access、Trunk和Hybrid这三种以太网端口所形成的链路又可归纳成两种以太网链路：接入链路（Access Link）和干道链路（Trunk Link）。它们是根据链路中允许通过的VLAN数据帧数量的不同来划分的。

● 接入链路（Access Link）：这是交换机直接连接用户主机的链路。通常情况下，主机并不需要知道自己属于哪个VLAN，主机硬件通常也不能识别带有VLAN标签的帧，所以主机通过接入链路发送和接收的数据帧都是Untagged帧。但一定要注意，接入链路不一定只允许来自一个VLAN的数据帧通过，只是Access端口链路才仅允许一个VLAN数据帧通过，在连接用户主机的Hybrid端口链路上同样允许来自多个VLAN的数据帧（不带标签）通过。

● 干道链路（Trunk Link）：这是用于交换机间的互连或交换机与路由器之间连接的链路。干道链路可以承载多个不同VLAN数据，数据帧在干道链路传输时，干道链路的两端设备需要识别数据帧属于哪个VLAN，所以在干道链路上传输的都是Tagged帧，除了该链路的PVID所属VLAN的帧（缺省为VLAN1）。

图6-6显示了以上两种链路类型及所传输的帧类型（带有Tag的帧和Untagged帧），从中可以发现，在交

换机设备之间的链路都属于干道链路，传输的是带有Tag的帧；而在交换机与主机设备之间的可以是接入链路，也可以是干道链路，具体要视主机所连接的交换机端口类型而定，但传输的都是Untagged的帧。

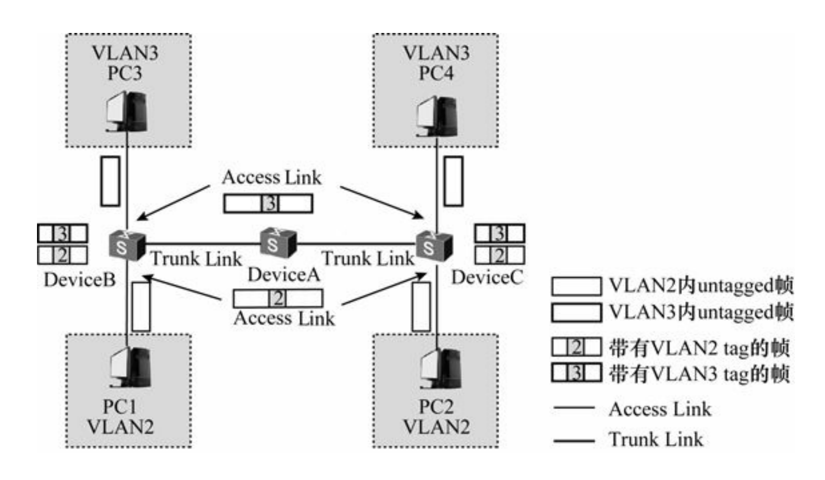


图6-6 两种链路类型及可传输的帧类型

说明

交换机在接收到帧后，会根据对应端口类型采取相应的数据收、发处理。如果帧需要通过另一台交换机转发，则该帧必须通过干道链路透传到对端交换设备上。为了保证其他交换设备能够正确处理帧中的VLAN信息，在干道链路上传输的帧必须打上VLAN标签。

当交换机最终确定帧出端口后，在将帧发送给主机前需要将 VLAN 标签从帧中删除，这样主机接收到的帧都是不带VLAN标签的以太网帧，也只有这样主机才可能识别。所以一般情况下，干道链路上传输的都是带VLAN标签的帧，接入链路上传输的都是不带 VLAN 标签的帧。这样处理的好处是网络中配置的VLAN 信息可以被所有交换设备正确处理，而主机不需要了解VLAN信息。

6.2.3 配置基于端口划分VLAN

基于端口划分VLAN是最简单，也是最常用的一种VLAN划分方式。它的划分思想就是通过把连接用户计算机的交换机端口指定到具体的 VLAN（是“交换机端口”与“VLAN“间的映射，不考虑用户计算机上任何配置）中来实现用户计算机的VLAN加入。

基于端口划分 VLAN 方式是一种静态 VLAN 划分方式，因为在这种划分方式中，只要把一个交换机端口划分到一个VLAN中，则所有连接在这个交换机端口的用户计算机都将成为该VLAN成员，也就是对具体的交换机端口来说，所连接的用户计算机是属于固定的VLAN。

1．基本配置思路

基于端口划分VLAN主要包括以下三项配置任务。

（1）创建所需的VLAN：如果VLAN已创建好，则此可直接略过。

（2）配置端口类型：基于端口划分VLAN时可以加入Access、Trunk和Hybrid这三种二层以太网端口。但要注意的是，在华为交换机中，二层以太网端口缺省情况下是Hybrid类型的，并且以不带标签方式加入VLAN 1。当然可以通过命令修改。但要注意表6-2中所介绍的这三种二层以太网端口的用途及对数据帧的收、发处理规则。

（3）把端口加入VLAN中：这是把用户计算机连接的交换机二层以太网端口加入你所期望并且已创建好的VLAN中。

2．配置步骤

基于端口划分VLAN的具体配置步骤如表6-3所示。

表6-3 基于端口划分VLAN的配置步骤

<table>
<tr><th>配置任务</th><th>步骤</th><th>命令</th><th>说明</th></tr>
<tr><td rowspan="2">创建并进入 VLAN 视图</td><td>1</td><td>system-view
例如：<HUAWEI> system-view</td><td>进入系统视图</td></tr>
<tr><td>2</td><td>vlan vlan-id
例如：[HUAWEI] vlan 2</td><td>创建 VLAN 并进入 VLAN 视图。如果 VLAN 已经创建，则直接进入 VLAN 视图。参数 vlan-id 的取值范围是 1～4 094。可用 undo vlan vlan-id 命令删除指定的 VLAN，但是 VLAN 1 是系统自带的 VLAN，不需要创建，也不可以删除</td></tr>
</table>

（续表）

<table>
<tr><th>配置任务</th><th>步骤</th><th>命令</th><th>说明</th></tr>
<tr><td rowspan="2">创建并进入 VLAN 视图</td><td>2</td><td>vlan vlan-id
例如：[HUAWEI] vlan 2</td><td>【说明】如果要同时创建多个 VLAN，则可使用 vlan batch { vlan-id1 [to vlan-id2] } &<1-10>命令，vlan-id1 表示第一个 VLAN 的编号，vlan-id2 表示最后一个 VLAN 的编号，取值范围都是 1～4 094 的整数，但 vlan-id2 的取值必须大于等于 vlan-id1，它与 vlan-id1 共同确定一个 VLAN 范围。如果不指定 vlan-id2 参数，则只创建 vlan-id1 所指定的 VLAN。可用 undo vlan batch { vlan-id1 [to vlan-id2] } &<1-10>命令删除指定的 VLAN</td></tr>
<tr><td>3</td><td>quit
例如：[HUAWEI-vlan2] quit</td><td>退出 VLAN 视图，返回系统视图</td></tr>
<tr><td rowspan="2">配置端口类型</td><td>4</td><td>interface interface-type interface-number
例如: [HUAWEI] interface gigabitethernet 0/0/1</td><td>键入要加入 VLAN 的二层以太网端口的接口类型和接口编号（注意：可以是 Eth-Trunk 口）。接口类型和接口编号之间可以输入空格，也可以不输入空格</td></tr>
<tr><td>5</td><td>port link-type { access | hybrid | runk }
例如：[HUAWEI-GigabitEthernet0/0/1] port link-type access</td><td>配置以上二层以太网端口的类型。命令中的选项说明如下。
（1）access：多选一选项，配置以上端口为 Access 类型
（2）hybrid：多选一选项，配置以上端口为 Hybrid 类型，这是缺省二层以太网端口类型
（3）trunk：多选一选项，配置以上端口为 Trunk 类型
在改变端口类型前，需要删除原端口类型下 VLAN 配置，即恢复接口只加入 VLAN1 的缺省配置。但本命令不可用于已经加入 Eth-Trunk 的物理接口，且在同一接口视图下多次使用本命令配置链路类型后，按最后一次配置生效
缺省情况下，S 系列交换机的以太网端口为 Hybrid 类型，可用 undo port link-type 命令恢复接口为缺省的链路类型</td></tr>
<tr><td>把以上 Access 端口加入到一个指定的 VLAN 中</td><td>6</td><td>port default vlan vlan-id
例如：[HUAWEI-GigabitEthernet0/0/1] port default vlan 2</td><td>（可选）将 Access 端口加入到指定的 VLAN 中，参数 vlan-id 的取值范围是 1～4 094 的整数。如果需要批量将端口加入 VLAN，可在 VLAN 视图下执行命令 port interface-type { interface- number1 [to interface-number2] } &<1-10>向 VLAN 中添加一个或一组端口
缺省情况下，所有交换机端口都以带标签方式的 Hybrid 类型加入 VLAN 1 中，可用 undo port default vlan 命令恢复该端口的缺省 VLAN 配置</td></tr>
</table>

（续表）

配置任务	步骤	命令	说明
把以上 Trunk 端口加入到一个或多个指定的 VLAN 中，并可选为以上 Trunk 端口配置缺省 VLAN	7	**port trunk allow-pass vlan** { *vlan-id1* [**to** *vlan-id2*] } <1-10> | **all** } 例如：[HUAWEI-GigabitEthernet0/0/1] **port trunk allow-pass vlan** 2 **to** 10	（可选）将以上 Trunk 端口加入到指定的 VLAN 中。命令中的参数和选项说明如下。 （1）*vlan-id1*：指定第一个 VLAN 的 ID 号，取值范围是 1～4 094 的整数 （2）**to** *vlan-id2*：可选参数，指定最后一个 VLAN 的 ID 号，取值范围为 1～4 094 的整数，与 *vlan-id1* 共同构成一个二选一参数 （3）<1-10>：表示前面的参数对最多可以重复 10 次，各段之间以空格分隔 （4）**all**：二选一选项，指定 Trunk 接口加入所有 VLAN 缺省情况下，Trunk 类型接口只加入了 VLAN 1，可用 **undo port trunk allow-pass vlan** { { *vlan-id1* [**to** *vlan-id2*] }<1-10> | **all** } 命令删除对应 Trunk 类型端口加入的指定 VLAN
	8	**port trunk pvid vlan** *vlan-id* 例如：[HUAWEI-GigabitEthernet0/0/1] **port trunk pvid vlan** 10	（可选）配置以上 Trunk 类型端口的缺省 VLAN，参数 *vlan-id* 的取值范围是 1～4 094 的整数 缺省情况下，Trunk 类型接口的缺省 VLAN 为 VLAN 1，可用 **undo port trunk pvid vlan** 命令恢复以上 Trunk 类型端口的缺省 VLAN（即 VLAN1） **【说明】**使用本命令配置 Trunk 类型端口缺省 VLAN 前，该 VLAN 必须已创建。但是，**缺省 VLAN 不一定是接口允许通过的 VLAN**，只有使用上一步介绍的 **port trunk allow-pass vlan** { { *vlan-id1* [**to** *vlan-id2*] } <1-10> | **all** }命令将缺省 VLAN 加入到允许的 VLAN 列表中才能转发缺省 VLAN 的数据帧
把以上 Hybrid 端口加入一个或多个指定的 VLAN 中，并可选为以上 Hybrid 端口配置缺省 VLAN	9	**port hybrid Untagged vlan** { { *vlan-id1* [**to** *vlan-id2*] } <1-10> | **all** } 例如：[HUAWEI-GigabitEthernet0/0/1] **port hybrid Untagged vlan** 2 **to** 10	（可选）将以上**连接用户主机**的 Hybrid 端口以 Untagged 方式加入指定的 VLAN 中，即指定的这些 VLAN 帧将以 Untagged 方式（去掉帧中原来的 VLAN 标签）通过接口向外（即向对端设备发送，不是向本地交换机内部发送）发送。参数同前面介绍的 **port trunk allow-pass vlan** { { *vlan-id1* [**to** *vlan-id2*] } <1-10> | **all** }命令的对应参数 **本命令仅对通过该端口向对端设备发送 VLAN 帧时起作用，不对接收的帧起作用** 缺省情况下，Hybrid 端口以 Untagged 方式加入 VLAN1。可用 **undo port hybrid vlan** { { *vlan-id1* [**to** *vlan-id2*] }<1-10> | **all** }命令删除以上 Hybrid 类型端口加入的指定 VLAN

（续表）

配置任务	步骤	命令	说明
把以上 Hybrid 端口加入一个或多个指定的 VLAN 中，并可选为以上 Hybrid 端口配置缺省 VLAN	10	**port hybrid Tagged vlan** { { *vlan-id1* [**to** *vlan-id2*] } &<1-10> | **all** } 例如：[HUAWEI-GigabitEthernet0/0/1] **port hybrid Tagged vlan** 2 **to** 10	（可选）将以上**连接网络设备**的 Hybrid 端口以 Tagged 方式加入指定的 VLAN 中，即指定的这些 VLAN 帧将以 Tagged 方式（保留帧中原来的 VLAN 标签）通过接口向外（即向对端设备发送，不是向本地交换机内部发送）发送。参数同前面介绍的 **port trunk allow-pass vlan** { { *vlan-id1* [**to** *vlan-id2*] } &<1-10> | **all** }命令的对应参数 **本命令仅对通过该端口向对端设备发送 VLAN 帧时起作用，不对接收的帧起作用** 缺省情况下，Hybrid 端口以 Untagged 方式加入 VLAN1，可用 **undo port hybrid vlan** { { *vlan-id1* [**to** *vlan-id2*] }&<1-10> | **all** }命令删除以上 Hybrid 类型端口加入的指定 VLAN
	11	**port hybrid pvid vlan** *vlan-id* 例如：[HUAWEI-GigabitEthernet0/0/1] **port hybrid pvid vlan** 2	（可选）配置以上 Hybrid 类型端口的缺省 VLAN ID，参数和注意事项同前面介绍的 **port trunk pvid vlan** *vlan-id* 命令的参数和注意事项 缺省情况下，所有接口的缺省 VLAN ID 为 VLAN 1，可用 **undo port hybrid pvid vlan** 命令恢复以上 Hybrid 类型端口的缺省 VLAN ID（即 VLAN1）

【经验之谈】 交换机端口从对端设备收到的帧有可能是Untagged的，如连接用户计算机时，但所有以太网帧在交换机内部都是以Tagged的形式进行处理、转发的，因此交换机必须给端口收到的Untagged帧加上VLAN标签。为了实现此目的，必须配置接口的缺省 VLAN，即 PVID。当该接口收到 Untagged 帧时给它加上该端口缺省 VLAN 的VLAN标签。当然如果是Access类型端口，不需要另外配置PVID，因为其PVID就是该端口唯一加入VLAN的VLAN ID。

【示例 1】配置GE0/0/1端口的链路类型为Trunk，允许VLAN 10～VLAN 30通过，并配置其缺省VLAN为VLAN 5。

```
<HUAWEI>system-view
[HUAWEI] interface gigabitethernet 0/0/1
[HUAWEI-GigabitEthernet0/0/1] port link-type trunk
[HUAWEI-GigabitEthernet0/0/1] port trunk allow-pass vlan 10 to 30
[HUAWEI-GigabitEthernet0/0/1] port trunk pvid vlan 5
```

【示例 2】配置GE0/0/1端口的链路类型为Hybrid，并以带标签方式加入VLAN 10～VLAN 30，并配置其缺省VLAN为VLAN5。

```
<HUAWEI>system-view
[HUAWEI] interface gigabitethernet 0/0/1
[HUAWEI-GigabitEthernet0/0/1] port link-type hybrid
[HUAWEI-GigabitEthernet0/0/1] port hybrid Tagged vlan 10 to 30
[HUAWEI-GigabitEthernet0/0/1] port hybrid pvid vlan 5
```

【示例 3】配置GE0/0/1端口的链路类型为Hybrid，并以不带标签方式加入VLAN 10和VLAN 30，并配置其缺省VLAN为VLAN5。

```
<HUAWEI>system-view
[HUAWEI] interface gigabitethernet 0/0/1
[HUAWEI-GigabitEthernet0/0/1] port link-type hybrid
[HUAWEI-GigabitEthernet0/0/1] port hybrid Untagged vlan 10 30
[HUAWEI-GigabitEthernet0/0/1] port hybrid pvid vlan 5
```

6.2.4 基于端口划分VLAN的配置示例

图 6-7 所示为一个小型企业局域网示例，拓扑结构中的两台交换机（SwitchA 和SwitchB）上各连接了许多进行不同业务操作的用户。现要把连接在SwitchA上的User1和连接在SwitchB上的User2都划分到VLAN 2中，而把连接在SwitchA上的User3和连接在SwitchB上的User4都划分到VLAN 3中。

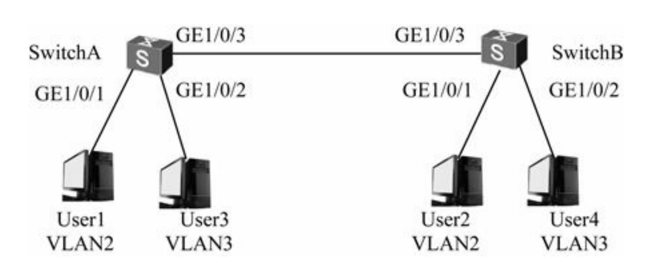


图6-7 基于接口划分VLAN示例拓扑结构

根据本章6.2.1节的介绍已经知道，用户PC机连接的端口既可以是Access类型的，又可以是不带标签的Hybrid类型的；而交换机之间连接的端口类型可以是Trunk类型，又可以是带标签的 Hybrid 类型。鉴于以上分析，本示例实际上可以有以下 4种配置方案。

方案一：用户端口采用Access类型，交换机间连接端口采用Trunk类型

（1）在SwitchA创建VLAN2和VLAN3，并将连接用户的端口类型都设置为Access类型，然后分别加入对应的VLAN中。SwitchB配置与SwitchA类似，不再赘述。

```
<HUAWEI>system-view
[HUAWEI] sysname SwitchA
[SwitchA] vlan batch 2 3    !---批量创建VLAN 2和VLAN 3
[SwitchA] interface gigabitethernet 1/0/1
[SwitchA-GigabitEthernet1/0/1] port link-type access    !---设置gigabitethernet1/0/1接口为Access类型
[SwitchA-GigabitEthernet1/0/1] port default vlan 2    !---把gigabitethernet1/0/1接口加入到VLAN 2中
[SwitchA-GigabitEthernet1/0/1] quit
[SwitchA] interface gigabitethernet 1/0/2
[SwitchA-GigabitEthernet1/0/2] port link-type access
[SwitchA-GigabitEthernet1/0/2] port default vlan 3
[SwitchA-GigabitEthernet1/0/2] quit
```

（2）配置SwitchA与SwitchB连接的端口类型为Trunk，同时允许VLAN 2和VLAN 3通过。SwitchB配置与SwitchA类似，不再赘述。

```
[SwitchA] interfacegigabitethernet 1/0/3
[SwitchA-GigabitEthernet1/0/3] port link-type trunk
[SwitchA-GigabitEthernet1/0/3] port trunk allow-pass vlan 2 to 3
```

方案二：用户端口采用 Access 类型，交换机间连接端口采用带标签的 Hybrid类型。

（1）把用户加入对应的VLAN中，配置方法同方案一中的第（1）步配置。

（2）配置SwitchA与SwitchB连接的端口类型为Hybrid，并以Tagged（带标签）方式同时加入VLAN 2和VLAN 3中。SwitchB配置与SwitchA类似，不再赘述。

```
[SwitchA] interfacegigabitethernet 1/0/3
```

[SwitchA-GigabitEthernet1/0/3] port link-type hybrid

[SwitchA-GigabitEthernet1/0/3] port hybrid Tagged vlan 2 to 3

方案三：用户端口采用不带标签的 Hybrid 类型，交换机间连接端口采用 Trunk类型。

（1）在SwitchA创建VLAN2和VLAN3，并将连接用户的端口类型都设置为Hybrid类型，然后分别以Untagged方式加入对应的VLAN中，并且把对应的VLAN ID设置这些Hybrid端口的PVID。SwitchB配置与SwitchA类似，不再赘述。

<HUAWEI>system-view

[HUAWEI] sysname SwitchA

[SwitchA] vlan batch 2 3

[SwitchA] interface gigabitethernet 1/0/1

[SwitchA-GigabitEthernet1/0/1] port link-type hybrid   !---设置gigabitethernet1/0/1接口为Hybrid类型

[SwitchA-GigabitEthernet1/0/1] port hybrid Untagged vlan 2   !---把gigabitethernet 1/0/1接口以Untagged方式加入到VLAN 2中

[SwitchA-GigabitEthernet1/0/1] port hybrid pvid vlan 2   !---设置gigabitethernet1/0/1接口的PVID为VLAN 2

[SwitchA-GigabitEthernet1/0/1] quit

[SwitchA] interface gigabitethernet 1/0/2

[SwitchA-GigabitEthernet1/0/2] port link-type hybrid

[SwitchA-GigabitEthernet1/0/2] port hybrid Untagged vlan 3

[SwitchA-GigabitEthernet1/0/2] port hybrid pvid vlan 3

[SwitchA-GigabitEthernet1/0/2] quit

（2）交换机间连接端口配置，配置方法同方案一中的第（2）步配置。

方案四：用户端口采用不带标签的 Hybrid 类型，交换机间连接端口采用带标签的Hybrid类型。

（1）把用户加入对应的VLAN中，配置方法同方案三中的第（1）步配置。

（2）交换机间连接端口配置，配置方法同方案二中的第（2）步配置。

验证配置结果如下。

将User1和User2配置在一个网段，比如192.168.100.0/24；将User3和User4配置在一个网段，比如192.168.200.0/24。经测试，证明User1和User2能够互相ping通，但是均不能ping通User3和User4。User3和User4能够互相ping通，但是均不能ping通User1和User2。由此可确认配置是正确并成功的。

6.3 基于MAC地址划分VLAN

基于MAC地址的VLAN划分方式是一种动态VLAN划分方式。它的划分思想是把用户计算机网卡上的MAC地址配置与某个VLAN进行关联（是“用户计算机网卡MAC地址”与“VLAN”之间的映射，不考虑用户计算机所连接的交换机端口），这样就可以实现无论该用户计算机连接在哪台交换机的二层以太网端口上都将保持所属的 VLAN不变。

也可以这么理解：基于MAC地址划分VLAN可以使无论用户计算机连接在哪 台交换机，也无论是连接在哪个交换机端口上，对应交换机端口都将成为该用户计算机网卡MAC地址所映射的VLAN的成员，而不需要在用户计算机改变所连接的端口时重新划分VLAN。这样就可以进一步提高终端用户的安全性（不会轻易被非法改变所属VLAN配置）和接入的灵活性（用户计算机可以在网络中根据实际需要随意移动）。

注意

基于MAC地址的VLAN划分方式只能在Hybrid交换机端口上进行，不能对其他类型端口上连接的用户计算机采用这种VLAN划分方式。

6.3.1 配置基于MAC地址划分VLAN

基于MAC地址划分的VLAN只处理Untagged数据帧，因为这里所介绍的VLAN都是单层VLAN标签的（QinQ可以实现双层VLAN标签），只有收到的数据帧中原来没有VLAN标签才可以根据交换机上所配置的MAC地址与VLAN ID映射关系，在数据帧中添加对应的VLAN标签。另外，基于MAC地址划分VLAN仅可在Hybrid端口上进行。这样一来就可使得基于MAC地址划分VLAN主要是针对终端用户设备，而非针对其他网络设备，因为在其他网络设备间连接的端口上发送的数据帧通常都是带有VLAN标签的，即使是Hybrid类型端口。

当交换机 Hybrid 端口收到的数据帧为 Untagged 数据帧时，端口会以数据帧的源MAC地址为根据去匹配MAC-VLAN映射表项。如果匹配成功，则在对应的数据帧中添加所匹配到的VLAN ID标签，然后按照对应的VLAN ID和优先级进行转发；如果匹配失败，则按其他匹配原则（如其他VLAN划分规则）进行匹配。当交换机端口收到的是Tagged 数据帧（仅在设备间连接的端口上才有可能），其处理方式和基于端口的 VLAN一样，根据Hybrid类型端口的数据收、发规则进行，参见表6-2。

1．基本配置思路

基于MAC地址划分VLAN的配置思路如下。

（1）创建要用于与用户主机MAC地址关联的VLAN。

（2）在以上创建的VLAN视图下关联用户MAC地址，建立MAC地址与VLAN的映射表，以确定哪些用户MAC地址可划分到以上创建的VLAN中。

（3）配置各用户连接的交换机二层以太网端口类型为Hybrid，并允许前面创建的基于 MAC 地址划分的 VLAN以不带 VLAN标签方式通过当前端口。因为华为交换机的所有二层以太网端口缺省都是 Hybrid 类型，所以缺省情况下，端口类型是不用配置的。

（4（）可选）配置VLAN划分方式的优先级，确保优先基于MAC地址划分VLAN。缺省情况下是优先基于 MAC地址划分 VLAN，但是可通过配置改变优先划分的方式。

（5）在Hybrid交换机端口上（注意，不一定要在连接用户计算机的**Hybrid**端口上配置 ）使能基于MAC地址划分VLAN功能，完成基于MAC地址划分VLAN。

2．配置步骤

基于MAC地址划分VLAN的配置步骤如表6-4所示。

表6-4 基于MAC地址划分VLAN的配置步骤

配置任务	步骤	命令	说明
创建并进入 VLAN 视图	1	**system-view** 例如：<HUAWEI> **system-view**	进入系统视图
	2	**vlan** *vlan-id* 例如：[HUAWEI] **vlan** 2	创建 VLAN 并进入 VLAN 视图。如果 VLAN 已经创建，则直接进入 VLAN 视图。其他说明参见表 6-3 中的第 2 步
配置 MAC 地址与 VLAN 映射表项	3	**mac-vlan mac-address** *mac-address* [*mac-address-mask* | *mac-address-mask-length*] [**priority** *priority*] 例如：[HUAWEI-vlan2] **mac-vlan mac-address 22-33-44**	创建用户计算机网卡 MAC 地址与 VLAN 的映射表项。命令中的参数说明如下： （1）*mac-address*：指定要与 VLAN 关联的用户计算机 MAC 地址，格式为 H-H-H，其中 H 为 4 位的十六进制数，可以输入 1～4 位，如 00e0、fc01，但这里的 MAC 地址不可设置为全 F、全 0 或组播 MAC 地址 （2）*mac-address-mask*：二选一可选参数，指定以上 MAC 地址的掩码，格式为 H-H-H，其中 H 为 1 至 4 位的十六进制数。MAC 地址掩码是用来确定在创建 MAC 地址与 VLAN 映射表项时对 MAC 地址匹配的比特位，只有值为 1 的比特位才进行匹配。如果要精确匹配一个 MAC 地址，则 MAC 地址掩码为 FFFF-FFFF-FFFF，但如果想要为一批具有某些相同特点的 MAC 地址创建与 VLAN 的映射表项，则其掩码不能是 FFFF-FFFF-FFFF，可以是像 FFFF-FFFF-0000 这样的，这样只要前 32 位是一样的 MAC 地址都要创建与 VLAN 的映射表项。**但 S7700、S9300 和 S9700 系列不支持该可选参数** （3）*mac-address-mask-length*：二选一可选参数，指定 MAC 地址掩码长度，整数形式，取值范围是 1～48。**但 S7700、S9300 和 S9700 系列不支持该可选参数** （4）**priority** *priority*：可选参数，指定以上 MAC 地址所对应的 VLAN 的 802.1p 优先级。取值范围是 0～7，值越大优先级越高。缺省值是 0。配置过程中，可以指定 MAC 地址对应的 VLAN 的 802.1p 优先级，用于当交换机阻塞时，优先发送优先级高的数据包。**但 S2700 系列不支持该可选参数** 缺省情况下，MAC 地址与 VLAN 没有关联，可用 **undo mac-vlan mac-address** { **all** | *mac-address* [*mac-address-mask* | *mac-address-mask-length*] }命令取消指定 MAC 地址与 VLAN 关联
	如果有多个 MAC 地址与 VLAN 映射表项，则重复第 3 步。但要注意，**如果映射的 VLAN 不一样，则一定要在对应的 VLAN 视图下配置映射**		
	4	**quit** 例如：[HUAWEI-vlan2]	退去 VLAN 视图，返回系统视图
配置 Hybrid 端口属性并启用基于 MAC 地址 VLAN 划分功能	5	**interface** *interface-type interface-number* 例如：[HUAWEI] **interface gigabitethernet 0/0/1**	键入要采用基于 MAC 地址划分 VLAN 的交换机端口（**注意：可以是 Eth-Trunk 口，且包括但不限于连接用户计算机的端口**）的接口类型和接口编号。接口类型和接口编号之间可以输入空格也可以不输入空格

（续表）

配置任务	步骤	命令	说明
配置 Hybrid 端口属性并启用基于 MAC 地址 VLAN 划分功能	6	**port link-type hybrid** 例如：[HUAWEI-GigabitEthernet0/0/1] **port link-type hybrid**	（可选）配置以上二层以太网端口类型为 Hybrid 类型。但因为 Hybrid 类型是华为交换机二层以太网端口的缺省类型，故多数情况不需要配置 在改变端口类型前，需要删除原端口类型下 VLAN 配置，即恢复接口只加入 VLAN1 的缺省配置。**但本命令不可用于已经加入 Eth-Trunk 的物理接口，且在同一接口视图下多次使用本命令配置链路类型后，按最后一次配置生效** 缺省情况下，接口的链路类型为 Hybrid，可用 **undo port link-type** 命令恢复端口的链路类型为缺省的 Hybrid 类型
	7	**port hybrid Untagged vlan** { { *vlan-id1* [**to** *vlan-id2*] } &<1-10> | **all** } 例如：[HUAWEI-GigabitEthernet0/0/1] **port hybrid Untagged vlan** 2 **to** 10	配置以上 Hybrid 类型端口以 Untagged 方式加入指定的 VLAN 中，即指定这些 VLAN 帧将以 Untagged 方式（去掉帧中原来的 VLAN 标签）通过接口向外（即向对端设备发送，不是向本地交换机内部发送）发送出去。其他方面的说明参见表 6-3 中的第 5 步
	8	**vlan precedence mac-vlan** 例如：[HUAWEI-GigabitEthernet0/0/1] **vlan precedence mac-vlan**	（可选）指定优先基于 MAC 地址划分 VLAN。不过其实也可不用配置，因为缺省情况下也是优先基于 MAC 地址划分 VLAN。也可用 **undo vlan precedence** 命令恢复该配置为缺省的基于 MAC 地址划分 VLAN
	9	**mac-vlan enable** 例如：[HUAWEI-GigabitEthernet0/0/1] **mac-vlan enable**	在以上 Hybird 端口上（**通常是在网络设备之间连接的 Hybird 端口上集中配置，而不是为每个连接用户计算机的 Hybrid 端口上配置**）使能基于 MAC 地址划分 VLAN。当端口收到 Untagged 数据帧时会以数据帧的源 MAC 地址去匹配 MAC-VLAN 表项。如果匹配成功，则按照匹配到的 VLAN ID 进行转发；如果匹配失败，则按照优先级选择其他匹配原则继续进行匹配。而当收到 Tagged 数据帧时，则按照基于端口划分 VLAN 进行转发 缺省情况下，未使能基于 MAC 地址划分 VLAN 功能，可用 **undo mac-vlan enable** 命令取消该端口的 MAC VLAN 功能
	对其他需要采用基于 MAC 地址划分 VLAN 的 Hybrid 端口重复以上第 5～9 步		

注意

如果某VLAN配置为MAC VLAN，要删除该VLAN，必须先使用undo mac-vlan mac-address { all |mac-address [mac-address-mask |mac-address-mask-length] }命令删除所有MAC地址与VLAN的关联N。且在S5700EI上多次使用mac-vlan mac-address命令关联VLAN和MAC地址时，如果指定的mac-address相同，则指定了MAC地址掩码的配置优先生效。当一个MAC地址关联了MAC VLAN后，则不可以再用于配置其他MAC VLAN，以避免一个计算机用户加入多个VLAN中；多次使用mac-vlan mac-address命令把当前VLAN与不同MAC地址进行关联时，配置结果按多次累加生效，也就相当于创建了多个MAC地址与VLAN映射表项。但MAC-VLAN与MUX-VLAN冲突，不允许在同一接口上同时配置这两种VLAN，且MAC-VLAN对接收到的VLAN ID为 0的数据帧不生效。

【示例 1】配置MAC地址22-33-44与VLAN100关联。

<HUAWEI>system-view

[HUAWEI] vlan 100

[HUAWEI-vlan100] mac-vlan mac-address 22-33-44

【示例 2】配置GE0/0/1端口优先采用基于MAC地址划分VLAN，并使能基于MAC地址的VLAN划分功能。

<HUAWEI> system-view

[HUAWEI] interfacegigabitethernet0/0/1

[HUAWEI-GigabitEthernet0/0/1] vlan precedence mac-vlan

[HUAWEI-GigabitEthernet0/0/1] mac-vlan enable

6.3.2 基于MAC地址划分VLAN的配置示例

某个公司的基本网络结构如图6-8所示。为了提高部门内的信息安全，每个部门的员工划分到一个VLAN中。现假设在工程部中有PC1、PC2、PC3三个用户，现要求在该部门中仅这几台PC可以通过SwitchA、Switch访问公司网络，如换成其他PC则不能访问。

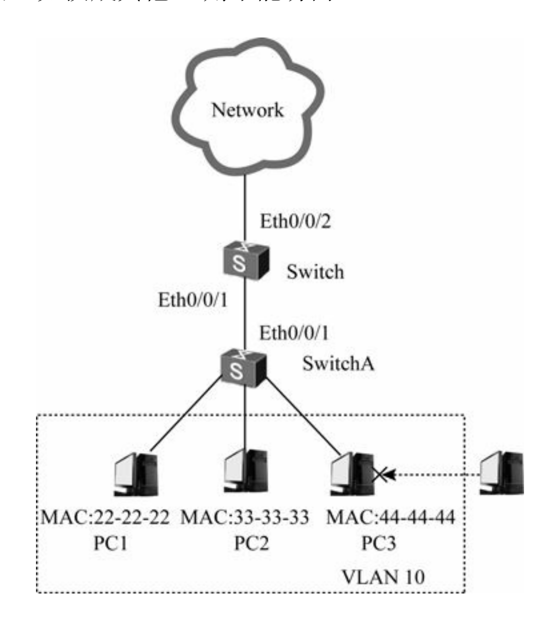


图6-8 基于MAC地址的VLAN划分配置示例拓扑结构

根据以上要求，可以针对工程部的以上三台PC 配置基于MAC 地址划分的VLAN 10，将它们的MAC地址与VLAN绑定，从而可以防止非法PC访问公司网络。

1．配置思路分析

面对这样一个基于 MAC 地址的 VLAN 划分示例，大家最容易想到的是直接到SwitchA交换机上进行配置。但其实际通常不是这样配置的，还有一种更为简便的方法，那就是直接在Switch上启用基于MAC地址的VLAN划分功能。

本示例的基本配置思想如下。

（1）因为华为交换机上所有二层以太网端口缺省都是 Hybrid 类型，并且发送数据帧时都是不带VLAN 标签的，故其实完全可以让SwitchA全部采用缺省配置（当然如果该交换机的缺省配置有改变，需要重新恢复其缺省配置），这样到达Switch交换机的数据帧都是不带VLAN标签的。

（2）然后通过在Switch交换机与SwitchA交换机连接的 Eth0/0/1 端口上配置不带标签发送特性的Hybrid类型，允许来自VLAN 10的数据帧通过，并且启用基于MAC地址划分VLAN功能，就可使得连接在SwitchA上的PC1、PC2和PC3发送的数据帧在到达 Switch后自动打 上对应的VLAN 10标签。

（3）最后将Switch交换机的Eth0/0/2端口配置为带标签的Hybrid类型，并允许VLAN 10的数据帧通过即可。

2．配置步骤

SwitchA交换机上全部采用缺省配置（所有二层以太网端口类型缺省为Hybrid，并且以Untagged方式加

入到VLAN1），所以无需另外配置（如果交换机上的缺省配置发生了改变，则需要先恢复到缺省配置）。现在只需要在 Switch 交换机上做如下配置。

（1）创建VLAN。

这里要创建 PC1、PC2 和 PC3 这三个 PC 用户要通过 MAC-VLAN 功能加入的VLAN 10。

<HUAWEI> system-view

[HUAWEI] vlan 10

（2）创建PC的MAC地址与VLAN 10关联。

[HUAWEI-Vlan10] mac-vlan mac-address 22-22-22

[HUAWEI-Vlan10] mac-vlan mac-address 33-33-33

[HUAWEI-Vlan10] mac-vlan mac-address 44-44-44

[HUAWEI-Vlan10] quit

（3）配置接口加入的VLAN。

[HUAWEI] interface ethernet 0/0/1

[HUAWEI-Ethernet0/0/1] port hybrid Untagged vlan 10　!---指定允许VLAN 10的数据帧通过，且发送时不带VLAN标签

[HUAWEI-Ethernet0/0/1] quit

[HUAWEI] interface ethernet 0/0/2

[HUAWEI-Ethernet0/0/2] port hybrid Tagged vlan 10　!---指定允许VLAN 10的数据帧通过，且发送时必须带有VLAN标签

[HUAWEI-Ethernet0/0/2] quit

（4）在连接SwitchA的Eth0/0/1端口上使能基于MAC地址划分VLAN功能。

[HUAWEI] interface ethernet 0/0/1

[HUAWEI-Ethernet0/0/1] mac-vlan enable

[HUAWEI-Ethernet0/0/1] quit

通过以上配置就可以实现PC1、PC2、PC3成功访问公司网络，并且连接在SwitchA端口上的计算机换成其他外来人员的PC时不能访问，因为在Switch交换机上并没有配置对应的MAC地址与VLAN映射表项，提高了网络安全性能。

6.4 基于子网划分VLAN

基于子网划分VLAN是基于数据帧中上层（网络层）IP地址或所属IP网段进行的VLAN划分，与下节将要介绍的“基于协议划分VLAN”统称为“基于网络层划分VLAN”，也属于动态 VLAN 划分方式，既可减少手工配置 VLAN 的工作量，又可保证用户自由地增加、移动和修改。基于子网划分VLAN适用于对安全性需求不高，对移动性和简易管理需求较高的场景中。

基于子网VLAN的划分思想是把用户计算机网卡上的IP地址配置与某个VLAN 进行关联（是“用户计算机网卡 IP 地址”与“VLAN”之间的映射，不考虑用户计算机所连接的交换机端口），这样与上节介绍的基于MAC地址划分VLAN一样，也可以实现无论该用户计算机连接在哪台交换机的二层以太网端口上都将保持所属的VLAN不变。

与上节介绍的基于MAC地址的VLAN划分一样，基于IP子网划分的VLAN也只处理Untagged数据帧（原因同6.3节介绍），所以也只能在**Hybird**类型端口上进行划分，对于Tagged数据帧处理方式和基于端口划

分的VLAN一样。

6.4.1 配置基于IP子网划分VLAN

基于 IP子网划分 VLAN的基本原理也与基于 MAC 地址划分 VLAN的原理类似，只是原来的 MAC 地址改成了 IP 地址，即当设备端口接收到 Untagged 数据帧时，设备根据数据帧的源 IP 地址或指定网段来确定数据帧所属的 VLAN，并在数据帧中添加对应的VLAN ID标签，然后将数据帧自动划分到指定VLAN中传输。

1．基本配置思路

基于IP子网划分VLAN的配置思路与6.3.1节介绍的基于MAC地址划分VLAN的配置思路基本一样，只是把匹配的MAC地址换成IP地址，具体如下。

（1）创建用于与用户主机MAC地址关联的VLAN。

（2）在以上创建的VLAN视图下关联用户IP地址，建立IP地址与VLAN的映射表，以确定哪些用户IP地址可划分到以上创建的VLAN中。

（3）配置各用户连接的交换机二层以太网端口类型为 Hybrid，并允许前面创建的基于IP地址划分的VLAN以不带VLAN标签方式通过当前端口。因为华为交换机的所有二层以太网端口缺省都是Hybrid类型，所以缺省情况下，端口类型是不用配置的。

（4）（可选）配置VLAN划分方式的优先级，确保优先基于IP地址划分VLAN。缺省情况下是优先基于 MAC地址划分 VLAN，但是可通过配置改变优先划分的方式。

（5）在Hybrid交换机端口上（注意，不一定要在连接用户计算机的**Hybrid**端口上 ）使能基于IP地址划分VLAN功能，完成基于IP地址划分VLAN。

2．配置步骤

基于IP地址划分VLAN的配置步骤与6.3.1节介绍的基于MAC地址划分VLAN的配置步骤基本一样，只是有些命令上的差异而已，具体如表6-5所示。

表6-5 基于IP子网划分VLAN的配置步骤

配置任务	步骤	命令	说明
创建并进入 VLAN 视图	1	**system-view** 例如：<HUAWEI> **system-view**	进入系统视图
	2	**vlan** *vlan-id* 例如：[HUAWEI] **vlan** 2	创建 VLAN 并进入 VLAN 视图。如果 VLAN 已经创建，则直接进入 VLAN 视图。其他说明参见表6-3 的第 2 步

（续表）

配置任务	步骤	命令	说明
配置 IP 地址与 VLAN 映射表项	3	**ip-subnet-vlan** [*ip-subnet-index*] **ip** *ip-address* { *mask* | *mask-length* } [**priority** *priority*] 例如：[HUAWEI-vlan2] **ip-subnet-vlan ip** 192.168.0.10 24	将以上创建的 VLAN 与用户计算机的 IP 地址进行关联，建立映射表项。命令中的参数说明如下。 （1）*ip-subnet-index*：可选参数，指定 IP 子网索引值，取值范围为 1～12 的整数。子网索引可由用户指定，也可由系统根据 IP 子网划分 VLAN 的顺序自动产生 （2）*ip-address*：指定基于 IP 子网划分 VLAN 依据的源 IP 地址或网络地址，为点分十进制格式 （3）*mask*：二选一参数，指定以上 IP 地址的子网掩码，为点分十进制格式 （4）*mask-length*：二选一参数，指定以上 IP 地址的子网掩码前缀长度，取值范围为 1～32 的整数 （5）**priority** *priority*：可选参数，指定以上 IP 地址或网段对应的 VLAN 的 802.1p 优先级。取值范围是 0～7，值越大优先级越高。缺省值是 0。配置过程中，可以指定 IP 地址或网段对应 VLAN 的 802.1p 优先级，用于当交换机阻塞时，优先发送优先级高的数据包。 缺省情况下，没有配置基于 IP 子网划分 VLAN，可用 **undo ip-subnet-vlan** { *ip-subnet-index* [**to** *ip-subnet-end*] | **all** }命令删除基于 IP 子网划分的指定 VLAN
	如果有多个 IP 地址与 VLAN 映射表项，则重复第 3 步。但要注意，**如果映射的 VLAN 不一样，则一定要在对应的 VLAN 视图下配置映射**		
	4	**quit** 例如：[HUAWEI-vlan2]	退去 VLAN 视图，返回系统视图
配置 Hybrid 端口属性并启用基于 IP 地址 VLAN 划分功能	5	**interface** *interface-type interface-number* 例如：[HUAWEI] **interface gigabitethernet 0/0/1**	键入要采用基于 IP 地址划分 VLAN 的交换机端口（**注意：可以是 Eth-Trunk 口，且包括但不限于连接用户计算机的端口**）的接口类型和接口编号。接口类型和接口编号之间可以输入空格也可以不输入空格
	6	**port link-type hybrid** 例如：[HUAWEI-GigabitEthernet0/0/1] **port link-type hybrid**	（可选）配置以上二层以太网端口类型为 Hybrid 类型。有关其他方面的说明参见表 6-4 的第 6 步
	7	**port hybrid Untagged vlan** { { *vlan-id1* [**to** *vlan-id2*] } &<1-10> | **all** } 例如：[HUAWEI-GigabitEthernet0/0/1] **port hybrid Untagged vlan** 2 **to** 10	配置以上 Hybrid 类型端口以 Untagged 方式加入指定的 VLAN 中，即指定这些 VLAN 帧将以 Untagged 方式（去掉帧中原来的 VLAN 标签）通过接口向外（即向对端设备发送，不是向本地交换机内部发送）发送出去。其他说明参见表 6-3 的第 5 步
	8	**vlan precedence ip-subnet-vlan** 例如：[HUAWEI-GigabitEthernet0/0/1] **vlan precedence ip-subnet-vlan**	（可选）指定优先基于 IP 地址划分 VLAN 缺省情况下是优先基于 MAC 地址划分 VLAN，可用 **undo vlan precedence** 命令恢复该配置为缺省的基于 MAC 地址划分 VLAN

（续表）

配置任务	步骤	命令	说明
配置 Hybrid 端口属性并启用基于 IP 地址 VLAN 划分功能	9	**ip-subnet-vlan enable** 例如：[HUAWEI-GigabitEthernet0/0/1] **ip-subnet-vlan enable**	在以上 Hybrid 端口上使能基于 IP 地址划分 VLAN。这样，当端口收到 Untagged 数据帧时会以数据帧的源 IP 地址去匹配 IP-VLAN 表项。如果匹配成功，则按照匹配到的 VLAN ID 进行转发；如果匹配失败，则按照优先级选择其他匹配原则继续进行匹配。而当收到 Tagged 数据帧时，则按照基于端口划分 VLAN 进行转发 缺省情况下，未使能基于 IP 地址划分 VLAN 功能，可用 **undo mac-vlan enable** 命令取消该端口的 MAC VLAN 功能
	对其他需要采用基于 IP 地址划分 VLAN 的 Hybrid 端口重复以上第 5～9 步		

【示例 1】将 10.10.10.0/24网段与VLAN 3进行关联，采用基于 IP子网的方式划分VLAN，使得源 IP地址在该网段的报文可以分发到VLAN 3中传输。

<HUAWEI>system-view

[HUAWEI] vlan 3

[HUAWEI-vlan3] ip-subnet-vlan ip 10.10.10.0 255.255.255.0

【示例 2】配置GE0/0/1端口优先采用基于IP子网的VLAN划分方式，并使能基于IP子网划分VLAN的功能。

```
<HUAWEI>system-view
[HUAWEI] interface gigabitethernet 0/0/1
[HUAWEI-GigabitEthernet0/0/1] vlan precedence ip-subnet-vlan
[HUAWEI-GigabitEthernet0/0/1] ip-subnet-vlan enable
```

6.4.2 基于IP子网划分VLAN配置示例

本示例拓扑结构如图6-9所示，假设该公司拥有多种业务，如IPTV、VoIP、Internet 等，而且使用每种业务的用户IP地址网段各不相同。为了便于管理，现需要将同一种类型业务划分到同一VLAN中，不同类型的业务划分到不同VLAN中，分别为VLAN 100、VLAN 200和VLAN 300。当Switch接收到这些业务数据帧时根据帧中封装的源 IP 地址网段的不同自动为这些帧添加对应的 VLAN ID标签，最终实现通过不同的VLAN ID分流到不同的远端服务器上以实现业务互通。

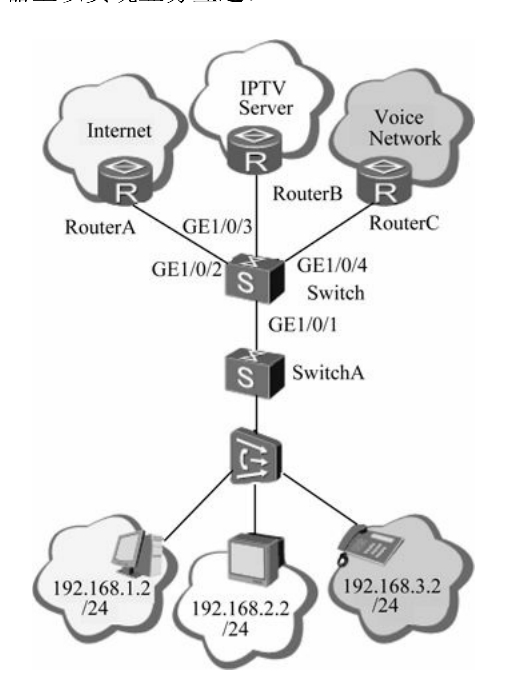


图6-9 基于IP子网的VLAN划分配置示例拓扑结构

1．配置思路分析

本示例其实与 6.3.2 节介绍的基于MAC地址划分VLAN的配置示例差不多，主要不同有两点：一是这是基于IP子网进行的VLAN划分，二是从Switch 上出去的数据帧要流向不同的服务器，这就需要在不同服务器所连接的交换机端口上配置仅允许某一个VLAN的数据帧通过。

同样，本示例也可以仅在 Switch上配置，使 Switch A上的配置全部保持缺省配置即可。

本示例的基本配置步骤如下。

（1）创建VLAN，确定每种业务所属的VLAN。

（2）关联IP子网和VLAN，实现根据数据帧中的源IP地址或指定网段确定VLAN。

（3）以正确的类型把各端口加入对应的VLAN，实现基于IP子网的VLAN通过当前端口。

（4）配置VLAN划分方式的优先级，确保优先选择基于IP子网划分VLAN。然后使能基于IP子网划分VLAN。

2．配置步骤

本示例在Switch上的具体配置步骤如下。

（1）为各业务用户创建所需的VLAN，即在Switch上创建VLAN100、VLAN200和VLAN300。

<HUAWEI>system-view

[HUAWEI] vlan batch 100 200 300

（2）关联IP子网与VLAN，并设置不同的优先级（其实优先级是可选配置）。

[HUAWEI] vlan 100

[HUAWEI-vlan100] ip-subnet-vlan 1 ip 192.168.1.2 24 priority 2 !---在 Switch 上配置 VLAN100 与 IP 地址192.168.1.2/24关联，优先级为2

[HUAWEI-vlan100] quit

[HUAWEI] vlan 200

[HUAWEI-vlan200] ip-subnet-vlan 1 ip 192.168.2.2 24 priority 3 !---在 Switch 上配置 VLAN200 与 IP 地址192.168.2.2/24关联，优先级为3

[HUAWEI-vlan200] quit

[HUAWEI] vlan 300

[HUAWEI-vlan300] ip-subnet-vlan 1 ip 192.168.3.2 24 priority 4 !---在 Switch 上配置 VLAN300 与 IP 地址192.168.3.2/24关联，优先级为4

[HUAWEI-vlan300] quit

（3）配置各端口类型及允许加入的 VLAN。注意，在启用基于 IP 子网划分 VLAN的 GE1/0/1 端口上要采用 Untagged 方式的 Hybrid 类型端口，并且要允许所有业务的VLAN数据帧通过；其他连接各数据服务器的端口可以是Trunk端口，也可以是Tagged方式的Hybrid类型端口（本示例仅以Trunk类型端口为例进行介绍），并且仅允许对应的VLAN数据帧通过。

[HUAWEI] interface gigabitethernet 1/0/1

[HUAWEI-GigabitEthernet1/0/1] port link-type hybrid

[HUAWEI-GigabitEthernet1/0/1] port hybrid Untagged vlan 100 200 300

[HUAWEI-GigabitEthernet1/0/1] quit

[HUAWEI] interface gigabitethernet 1/0/2

[HUAWEI-GigabitEthernet1/0/2] port link-type trunk

[HUAWEI-GigabitEthernet1/0/2] port trunk allow-pass vlan 100

[HUAWEI-GigabitEthernet1/0/2] quit

[HUAWEI] interface gigabitethernet 1/0/3

[HUAWEI-GigabitEthernet1/0/3] port link-type trunk

[HUAWEI-GigabitEthernet1/0/3] port trunk allow-pass vlan 200

[HUAWEI-GigabitEthernet1/0/3] quit

[HUAWEI] interface gigabitethernet 1/0/4

[HUAWEI-GigabitEthernet1/0/4] port link-type trunk

[HUAWEI-GigabitEthernet1/0/4] port trunk allow-pass vlan 300

[HUAWEI-GigabitEthernet1/0/4] quit

（4）在Switch上配置接口GE1/0/1优先采用基于IP子网进行VLAN划分，并使能基于IP子网划分VLAN功能。

[HUAWEI] interface gigabitethernet 1/0/1

[HUAWEI-GigabitEthernet1/0/1] vlan precedence ip-subnet-vlan

[HUAWEI-GigabitEthernet1/0/1] ip-subnet-vlan enable

[HUAWEI-GigabitEthernet1/0/1] quit

下面来验证以上配置结果，可在Switch上执行display ip-subnet-vlan vlan all命令查看基于IP子网划分的VLAN信息。从中可以看出，已按配置正确进行了VLAN划分。

[HUAWEI] display ip-subnet-vlan vlan all

Vlan	Index	IpAddress	SubnetMask	Priority
100	1	192.168.1.2	255.255.255.0	2
200	1	192.168.2.2	255.255.255.0	3
300	1	192.168.3.2	255.255.255.0	4

----

ip-subnet-vlan count: 3 total count: 3

6.5 基于协议划分VLAN

基于协议划分 VLAN 是指基于数据帧中的上层（网络层）协议类型进行的 VLAN划分。与前面介绍的基于MAC地址划分的VLAN和基于IP子网划分的VLAN一样，基于协议划分的VLAN也只处理Untagged数据帧，且也只能在**Hybrid**端口上进行配置，对于Tagged数据帧的处理方式和基于端口的VLAN一样。

基于协议 VLAN 的划分思想是把用户计算机上运行的网络层协议与某个 VLAN 进行关联（是"用户计算机网络层协议"与"VLAN"之间的映射，不考虑用户计算机所连接的交换机端口），这样也可以实现无论该用户计算机连接在哪台交换机的二层以太网端口上都将保持其所属的 VLAN 不变。启用基于协议划分 VLAN 功能后，当交换机端口接收到Untagged帧时，先识别帧的协议模板，然后确定数据帧所属的VLAN。如果端口配置了属于某些协议VLAN，且数据帧的协议模板匹配其中某个协议VLAN，则给数据帧打上该协议VLAN标签。如果端口原来配置了属于某些协议VLAN，但某次到达的数据帧的协议模板和所有协议VLAN都不匹配，则给数据帧打上端口PVID的VLAN标签（这点比较特殊，要充分注意 ）。

6.5.1 配置基于协议划分VLAN

基于协议划分VLAN与上节介绍的基于IP子网划分VLAN都属于基于网络层进行的VLAN划分，不同的是基于IP子网划分VLAN仅根据网络层中特定的IPv4协议中的IPv4地址或子网进行VLAN划分，而本节所介绍的基于协议划分VLAN是根据不同网 络层协议（包括IPv4、IPX、AppleTalk等协议）进行的VLAN划分，不是根据具体类型的网络层地址进行VLAN划分。

1．基本配置思路

因为基于协议划分VLAN是根据不同的网络层协议进行的，所以需要事先创建不同网络层协议与

VLAN 的映射表项，同时还要在交换机 Hybrid 端口上配置与对应的协议VLAN进行关联，以限定交换机端口仅可以加入特定的协议VLAN中。具体如下。

（1）创建各网络层协议所需关联的VLAN。

（2）在以上创建的VLAN视图下关联用户所用的网络层协议类型，建立网络层协议与VLAN的映射表，以确定哪些用户可划分到以上创建的VLAN中。

（3）配置各用户连接的交换机二层以太网端口类型为Hybrid，并允许前面创建的基于协议划分的VLAN 以不带 VLAN 标签方式通过当前端口。因为华为交换机的所有二层以太网端口缺省都是Hybrid类型，所以缺省情况下，端口类型是不用配置的。

（4）配置交换机 Hybrid端口与对应的协议 VLAN 进行关联。这样，当有关联的协议数据帧进入所关联的端口时，系统自动为该协议数据帧分配已经划分好的VLAN ID。

2．配置步骤

基于协议划分VLAN的具体配置步骤如表6-6所示。

表6-6 基于协议划分VLAN的配置步骤

<table>
<tr><th>配置任务</th><th>步骤</th><th>命令</th><th>说明</th></tr>
<tr><td rowspan="2">创建并进入 VLAN 视图</td><td>1</td><td>**system-view**
例如：<HUAWEI> **system-view**</td><td>进入系统视图</td></tr>
<tr><td>2</td><td>**vlan** *vlan-id*
例如：[HUAWEI] **vlan** 2</td><td>创建 VLAN 并进入 VLAN 视图。如果 VLAN 已经创建，则直接进入 VLAN 视图。其他说明参见表6-3 的第 2 步</td></tr>
<tr><td>配置网络层协议与 VLAN 映射表项</td><td>3</td><td>**protocol-vlan** [*protocol-index*] { **at** | **ipv4** | **ipv6** | **ipx** { **ethernetii** | **llc** | **raw** | **snap** } | **mode** { **ethernetii-etype** *etype-id1* | **llc dsap** *dsap-id* **ssap** *ssap-id* | **snap-etype** *etype-id2* } }
例如：[HUAWEI-vlan2] **protocol-vlan ipv4**</td><td>将以上创建的 VLAN 与特定的网络层协议进行关联。命令中的参数说明如下。
（1）*protocol-index*：可选参数，指定协议的索引值。如果不手工配置协议索引值，则系统会根据协议与VLAN 关联的先后顺序自动产生编号。除 S5700SI 的取值范围为 0～11 的整数外，其他支持基于协议划分 VLAN 功能的华为交换机的取值范围均为 0～15 的整数
（2）**at**：多选一选项，指定基于 AppleTalk 协议划分 VLAN
（3）**ipv4**：多选一选项，指定基于 IPv4 协议划分 VLAN
（4）**ipv6**：多选一选项，指定基于 IPv6 协议划分 VLAN
（5）**ipx**：多选一选项，指定基于 IPX 协议划分 VLAN
（6） **ethernetii**：多选一选项，指定 IPX 协议的以太网数据帧的封装格式为 Ethernet II 标准格式
（7）**llc**：多选一选项，指定 IPX 协议的以太网数据帧的封装格式为 802.3/802.2 LLC 标准格式
（8）**raw**：多选一选项，指定 IPX 协议的以太网数据帧的封装格式为 Ethernet 802.3 raw 标准格式
（9）**snap**：多选一选项，指定 IPX 协议的以太网数据帧的封装格式为 Ethernet 802.3 SAP 标准格式</td></tr>
</table>

（续表）

配置任务	步骤	命令	说明
配置网络层协议与VLAN映射表项	3	**protocol-vlan** [*protocol-index*] { **at** | **ipv4** | **ipv6** | **ipx** { **ethernetii** | **llc** | **raw** | **snap** } | **mode** { **ethernetii-etype** *etype-id1* | **llc dsap** *dsap-id* **ssap** *ssap-id* | **snap-etype** *etype-id2* } } 例如：[HUAWEI-vlan2] **protocol-vlan ipv4**	（10）**ethernetii-etype** *etype-id1*：多选一参数，指定匹配 Ethernet II 封装格式的协议类型值，取值范围是 600～ffff（除 800、809b、8137、86dd 以外的值）。 （11）**llc dsap** *dsap-id* **ssap** *ssap-id*：多选一参数，指定匹配 802.3/802.2 LLC 封装格式的目的服务访问点（DSAP）和源服务访问点（SSAP），取值范围均为 0～ff （12）**snap-etype** *etype-id2*：多选一参数，指定匹配 Ethernet 802.3 SAP 封装格式的协议类型值，取值范围是 600～ffff（除 800、809b、8137、86dd 以外的值） 缺省是没有建立任何网络层协议与 VLAN 关联的，可用 **undo protocol-vlan** { **all** | *protocol-index1* [**to** *protocol-index2*] }命令删除基于协议划分的指定 VLAN。二选一选项 **all** 用来指定删除所有基于协议划分的 VLAN，二选一参数 *protocol-index1* [**to** *protocol-index2*]用来指定要删除 VLAN 所对应的起始和终止协议索引值，取值范围是 0～15 的整数
	如果有多个网络层协议与 VLAN 映射表项，则重复第 3 步。但要注意，**如果映射的 VLAN 不一样，则一定要在对应的 VLAN 视图下配置映射**		
	4	**quit** 例如：[HUAWEI-vlan2]	退去 VLAN 视图，返回系统视图
配置 Hybrid 端口属性并与指定协议 VLAN 进行关联	5	**interface** *interface-type interface-number* 例如：[HUAWEI] **interface gigabitethernet 0/0/1**	键入要采用基于协议划分 VLAN 的交换机端口的接口类型和接口编号（**注意：可以是 Eth-Trunk 口**）。接口类型和接口编号之间可以输入空格也可以不输入空格
	6	**port link-type hybrid** 例如：[HUAWEI-GigabitEthernet0/0/1] **port link-type hybrid**	配置以上二层以太网端口类型为 Hybrid 类型。其他说明参见表 6-4 的第 6 步
	7	**port hybrid Untagged vlan** { { *vlan-id1* [**to** *vlan-id2*] } <1-10> | **all** } 例如：[HUAWEI-GigabitEthernet0/0/1] **port hybrid Untagged vlan** 2 **to** 10	配置以上 Hybrid 类型端口以 Untagged 方式加入指定的 VLAN 中，即指定这些 VLAN 帧将以 Untagged 方式（去掉帧中原来的 VLAN 标签）通过接口向外（即向对端设备发送，不是向本地交换机内部发送）发送出去。其他说明参见表 6-3 的第 5 步
	8	**protocol-vlan vlan** *vlan-id* { **all** | *protocol-index1* [**to** *protocol-index2*] } [**priority** *priority*] 例如：[HUAWEI-GigabitEthernet0/0/1] **protocol-vlan vlan 2 0**	把特定索引号的协议 VLAN 与特定交换机端口进行关联，以限定交换机端口可以加入的协议 VLAN，主要应用在根据不同协议类型采用不同传输路径的网络中。命令中的参数和选项说明如下。 （1）*vlan-id*：指定以上 Hybrid 端口要关联的协议 VLAN （2）**all**：二选一选项，指定要与所有协议索引值对应的，并由参数 *vlan-id* 指定的协议 VLAN 关联

（续表）

配置任务	步骤	命令	说明
配置 Hybrid 端口属性并与指定协议 VLAN 进行关联	8	**protocol-vlan vlan** *vlan-id* { **all** | *protocol-index1* [**to** *protocol-index2*] } [**priority** *priority*] 例如：[HUAWEI-GigabitEthernet0/0/1] **protocol-vlan vlan 2 0**	（3）*protocol-index1* [**to** *protocol-index2*]：二选一参数，指定仅与指定协议索引起始值和终止值范围内，由参数 *vlan-id* 指定的协议 VLAN 关联，取值范围均为 0～15 的整数。如果不手工配置协议索引值，则系统会根据协议与 VLAN 关联的先后顺序自动产生编号 （4）**priority** *priority*：可选参数，指定所关联的以上协议 VLAN 的 802.1p 优先级 可用 **undo protocol-vlan** { **all** | **vlan** *vlan-id* { **all** | *protocol-index1* [**to** *protocol-index2*] } }命令取消以上端口与指定协议 VLAN 的关联
	对其他需要采用基于协议划分 VLAN 的 Hybrid 端口重复以上第 5～8 步		

【示例 1】把 IPv4协议报文划分到VLAN 3中。

<HUAWEI>system-view

[HUAWEI] vlan 3

[HUAWEI-vlan3] protocol-vlan ipv4

【示例 2】配置GE0/0/1端口关联协议VLAN 2（协议的索引值为 0），即相当于把GE0/0/1端口加入协议VLAN 2中。

<HUAWEI>system-view

[HUAWEI] interface gigabitethernet 0/0/1

[HUAWEI-GigabitEthernet0/0/1] protocol-vlan vlan 2 0

6.5.2 基于协议划分VLAN的配置示例

本示例拓扑结构如图6-10所示。现假设该公司拥有多种业务，如IPTV、VoIP、Internet等，而且每种业务所采用的协议各不相同。为了便于管理，减少人工配置VLAN的工作 量，现需要将同一种类型业务划分到同一VLAN 中，不同类型的业务划分到不同VLAN中。本示例中，VLAN 10中的用户采用IPv4协议与远端用户通信，而VLAN 20中的用户采用IPv6协议与远端服务器通信，现要通过不同的VLAN ID分流到不同的远端服务器上以实现业务互通。

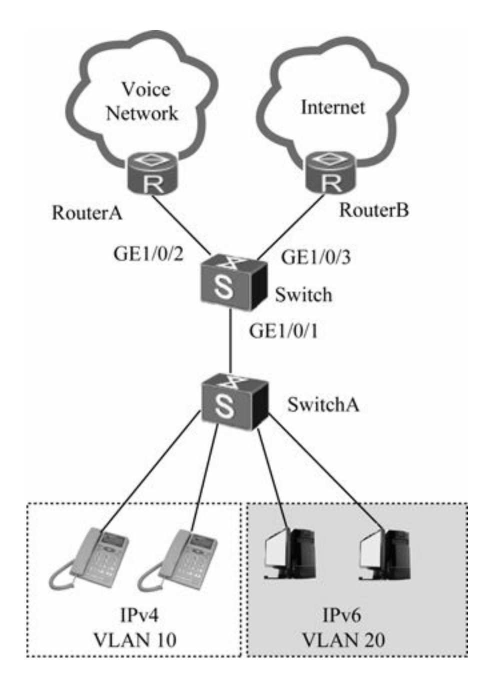


图6-10 基于协议划分VLAN配置示例拓扑结构

1．配置思路分析

本示例中需要创建两个协议 VLAN：VLAN 10和 VLAN 20，分别对应于 IPv4和 IPv6，所以事先要创建这两个 VLAN，然后分别与对应的协议进行关联。除此之外，还要在对应的Hybrid 端口上允许对应的协议VLAN通过，并与指定的协议VLAN进行关联。

与本章前面介绍的基于MAC地址划分 VLAN和基于IP子网划分VLAN的配置示例一样，本示例也仅需在Switch上配置，而保持SwitchA上全部为缺省配置。其基本配置思路如下。

（1）创建VLAN，确定每种业务所属的协议VLAN。

（2）关联协议和VLAN，实现根据端口接收到的数据帧所属的网络层协议类型给数据帧分配不同的VLAN ID。

（3）配置端口加入VLAN，并允许基于协议的VLAN通过当前端口。

（4）关联接口和对应的协议VLAN，使有关联的协议进入关联的接口时，系统自动为该协议分配已经划分好的VLAN ID。

2．配置步骤

本示例在Switch上的具体配置步骤如下。

（1）创建所需的协议VLAN 10和VLAN 20。

```
<HUAWEI>system-view
[HUAWEI] vlan batch 10 20
```

（2）配置网络层协议与以上协议VLAN的关联。

```
[HUAWEI] vlan 10
[HUAWEI-vlan10] protocol-vlan ipv4
[HUAWEI-vlan10] quit
[HUAWEI] vlan 20
[HUAWEI-vlan20] protocol-vlan ipv6
[HUAWEI-vlan20] quit
```

（3）配置端口类型及允许通过的协议VLAN。注意，与SwitchA连接的GE1/0/1端口要允许所有的协议VLAN 通过，并且必须是 Hybrid 类型；连接各业务服务器的交换机端口可以是带VLAN标签的Hybrid或Trunk端口类型，但仅允许对应的协议VLAN通过。

```
!----配置GE1/0/1端口为Hybrid类型，并同时允许VLAN 10和VLAN 20通过
[HUAWEI] interface gigabitethernet 1/0/1
[HUAWEI-GigabitEthernet1/0/1] port link-type hybrid
[HUAWEI-GigabitEthernet1/0/1] port hybrid Untagged vlan 10 20
[HUAWEI-GigabitEthernet1/0/1] quit
!----配置GE1/0/2端口为Trunk类型，但仅允许VLAN 10通过
[HUAWEI] interfacegigabitethernet 1/0/2
[HUAWEI-GigabitEthernet1/0/2] port link-type trunk
[HUAWEI-GigabitEthernet1/0/2] port trunk allow-pass vlan 10
[HUAWEI-GigabitEthernet1/0/2] quit
!----配置GE1/0/3端口为Trunk类型，但仅允许VLAN 20通过
[HUAWEI] interfacegigabitethernet 1/0/3
[HUAWEI-GigabitEthernet1/0/3] port link-type trunk
[HUAWEI-GigabitEthernet1/0/3] port trunk allow-pass vlan 20
[HUAWEI-GigabitEthernet1/0/3] quit
```

（4）配置GE1/0/1端口关联所需的协议VLAN，并为它们指定不同的优先级。

```
[HUAWEI] interface gigabitethernet 1/0/1
[HUAWEI-GigabitEthernet1/0/1] protocol-vlan vlan 10 all priority 5    !---配置GE1/0/1端口与VLAN10关联，优先级是5
[HUAWEI-GigabitEthernet1/0/1] protocol-vlan vlan20 all priority 6    !----配置GE1/0/1端口与VLAN20关联，优先级是6
[HUAWEI-GigabitEthernet1/0/1] quit
```

下面可以通过执行**display protocol-vlan interface all** 命令查看端口关联协议VLAN 的配置信息。从中可以看出，对应协议VLAN已成功配置。

```
<HUAWEI >display protocol-vlan interface all
-------------------------------------------------------------------------------
Interface    VLAN    Index    Protocol Type    Priority
-------------------------------------------------------------------------------
GigabitEthernet1/0/1    10    0    ipv4    5
GigabitEthernet1/0/1    20    0    ipv6    6
```

6.6 基于策略划分VLAN

基于策略划分VLAN也可称为Policy VLAN，是根据一定的策略进行VLAN划分的，可实现用户终端的即插即用功能，同时可为终端用户提供安全的数据隔离。这里的策略主要包括“基于MAC地址+IP地址”组合策略和“基于MAC地址+IP地址+端口”组合策略两种。

6.6.1 配置基于策略划分VLAN

基于策略划分VLAN是指在交换机上绑定终端的MAC地址、IP地址或交换机端口，并与VLAN关联，以证实只有符合条件的终端才能加入指定VLAN。符合策略的终端才可以加入指定的VLAN，相当于采用了IP地址与MAC地址双重绑定，甚至再加上与所连接的交换机端口的三重绑定，一旦配置就可以禁止用户修改IP地址或MAC地址，甚至禁止改变所连接的交换机端口，否则会导致终端从指定VLAN中退出，可能访问不了指定的网络资源。

与前面介绍的基于MAC地址、基于IP子网、基于协议划分的VLAN一样，基于策略划分的VLAN也只处理Untagged数据帧（所以也只能在Hybrid端口上进行配置），对于 Tagged 数据帧处理方式和基于端口划分的 VLAN 一样。当设备端口接收到Untagged数据帧时，设备根据用户数据帧中的“源MAC地址”和“源IP地址”字段值与交换机上配置的“MAC地址和IP地址”，或者“MAC地址和IP地址+交换机端口”组合策略来确定数据帧所属的VLAN，然后将数据帧自动划分到指定VLAN中传输。

1．基本配置思路

基于策略划分VLAN的基本配置思路比较简单，具体如下。

（1）创建各策略所需关联的VLAN。

（2）在以上创建的 VLAN 视图下关联不同的策略，建立特定策略与 VLAN 的映射表，以确定哪些用户可划分到以上创建的VLAN中。

（3）配置各用户连接的交换机二层以太网端口类型为 Hybrid，并允许前面创建的基于策略划分的 VLAN 以不带 VLAN 标签方式通过当前端口。因为华为交换机的所有二层以太网端口缺省都是Hybrid类型，所以缺省情况下，端口类型是不用配置的。

2．配置步骤

基于策略划分VLAN的具体配置步骤如表6-7所示。

表6-7 基于策略划分VLAN的配置步骤

配置任务	步骤	命令	说明
创建并进入 VLAN 视图	1	**system-view** 例如：<HUAWEI> **system-view**	进入系统视图
	2	**vlan** *vlan-id* 例如：[HUAWEI] **vlan** 2	创建 VLAN 并进入 VLAN 视图。如果 VLAN 已经创建，则直接进入 VLAN 视图。其他说明参见表 6-3 的第 2 步
配置策略与 VLAN 映射表项	3	**policy-vlan mac-address** *mac-address* **ip** *ip-address* [**interface** *interface-type interface-number*] [**priority** *priority*] 例如：[HUAWEI-vlan2] **policy-vlan mac-address 1-1-1 ip 10.10.10.1 priority 7**	将以上创建的 VLAN 与特定的策略进行关联。命令中的参数说明如下。 （1）**mac-address** *mac-address*：指定策略 VLAN 依据的源 MAC 地址，格式为 H-H-H。其中 H 为 4 位的十六进制数，可以输入 1～4 位，如 00e0、fc01。当输入不足 4 位时，表示前面的几位为 0，如输入 e0，等同于 00e0。MAC 地址不可设置为 0000-0000-0000、FFFF-FFFF-FFFF 和组播地址 （2）**ip** *ip-address*：指定策略 VLAN 依据的源 IP 地址，格式为点分十进制格式 （3）**interface** *interface-type interface-number*：可选参数，指定应用 MAC 地址和 IP 地址组合策略的交换机端口（**注意：可以是 Eth-Trunk 口**）。如果指定该参数，MAC 地址和 IP 地址组合策略只应用到指定 VLAN 中指定的端口上，否则 MAC 地址和 IP 地址组合策略将应用到指定 VLAN 中所有的端口上 （4）**priority** *priority*：可选参数，指定以上策略所对应的 VLAN 的 802.1p 优先级，取值范围为 0～7 的整数，值越大优先级越高。缺省值是 0 缺省情况下，没有配置基于策略划分 VLAN，可用 **undo policy-vlan** { **all** | **mac-address** *mac-address* **ip** *ip-address* [**interface** *interface-type interface-number*] } 命令删除基于策略划分的指定 VLAN。二选一选项 **all** 用来指定删除所有基于策略划分的 VLAN。如果要删除被设置为策略 VLAN 的 VLAN，需要先执行以上 **undo policy-vlan** 命令删除 Policy VLAN 后，才能够删除该 VLAN
	如果有多个策略与 VLAN 映射表项，则重复第 3 步。但要注意，**如果映射的 VLAN 不一样，则一定要在对应的 VLAN 视图下配置映射**		
	4	**quit** 例如：[HUAWEI-vlan2]	退去 VLAN 视图，返回系统视图
配置 Hybrid 端口属性并允许对应的策略 VLAN 通过	5	**interface** *interface-type interface-number* 例如：[HUAWEI] **interface gigabitethernet 0/0/1**	键入要采用基于策略划分 VLAN 的交换机端口的接口类型和接口编号（**注意：可以是 Eth-Trunk 口**）。接口类型和接口编号之间可以输入空格也可以不输入空格
	6	**port link-type hybrid** 例如：[HUAWEI-GigabitEthernet0/0/1] **port link-type hybrid**	配置以上二层以太网端口类型为 Hybrid 类型。其他说明参见表 6-4 的第 6 步

（续表）

配置任务	步骤	命令	说明
配置 Hybrid 端口属性并允许对应的策略 VLAN 通过	7	**port hybrid Untagged vlan** { { *vlan-id1* [**to** *vlan-id2*] } &<1-10> | **all** } 例如：[HUAWEI-GigabitEthernet0/0/1] **port hybrid Untagged vlan 2 to 10**	配置以上 Hybrid 类型端口以 Untagged 方式加入指定的 VLAN 中，即指定这些 VLAN 帧将以 Untagged 方式（去掉帧中原来的 VLAN 标签）通过接口向外（即向对端设备发送，不是向本地交换机内部发送）发送出去。其他说明参见表 6-3 的第 5 步
	对其他需要采用基于策略划分 VLAN 的 Hybrid 端口重复以上第 5～7 步		

【示例 】配置基于组合策略，把MAC地址为0-1-1，IP地址为1.1.1.1的主机划分到VLAN 2中，并配置该VLAN的 802.1p优先级是 7。

```
<HUAWEI> system-view
[HUAWEI] vlan 2
[HUAWEI-vlan2] policy-vlan mac-address 0-1-1 ip 1.1.1.1 priority 7
```

6.6.2 基于策略划分VLAN的配置示例

本示例拓扑结构如图6-11所示。现要把User1（MAC地址为1-1-1，IP地址为1.1.1.1）绑定在SwitchA的GE1/0/1端口上，把User2（MAC地址为2-2-2，IP地址为2.2.2.2）绑定在 SwitchB的GE1/0/1端口上，并把它们划分到VLAN 2中；把User3（MAC地址为3-3-3，IP地址为3.3.3.3）绑定在SwitchA的GE1/0/2端口上，把User4（MAC地址为4-4-4， IP地址为4.4.4.4）绑定在SwitchB的GE1/0/2端口上，并把它们划分到VLAN 3中。

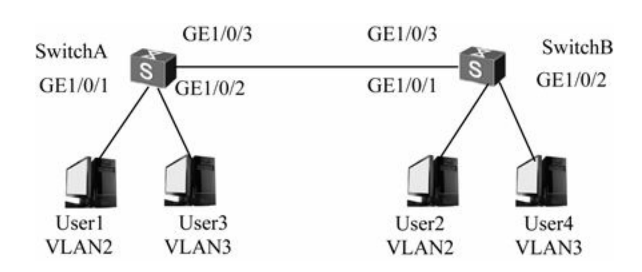


图6-11 基于策略划分VLAN配置示例拓扑结构

1．配置思路分析

基于策略划分VLAN的配置很简单，参照6.6.1节介绍的具体配置步骤可以得出本示例的以下3方面的基本配置任务。

（1）创建所需的策略VLAN。

（2）在对应的VLAN视图下配置基于用户计算机的MAC地址、IP地址的组合策略和应用策略的交换机端口。

（3）配置应用组合策略的Hybrid类型交换机端口允许所加入的VLAN通过。

2．配置步骤

通过以上配置思路分析后，下面的具体配置就比较简单了。

SwitchA上的配置：

（1）创建所需的策略协议VLAN 2和VLAN 3。

```
<HUAWEI>system-view
<HUAWEI>sysname SwitchA
[SwitchA] vlan batch 2 3
```

（2）配置MAC地址、IP地址和交换机端口组合策略与以上策略VLAN的关联，并为两个协议VLAN设置不同的802.1q的优先级值。

```
[SwitchA] vlan 2
[SwitchA -vlan2] policy-vlan mac-address 1-1-1 ip 1.1.1.1 gigabitEthernet1/0/1priority 7
[SwitchA -vlan2] quit
[SwitchA] vlan 3
[SwitchA -vlan20] policy-vlan mac-address 3-3-3 ip 3.3.3.3 gigabitEthernet1/0/2priority 5
[SwitchA-vlan20] quit
```

（3）配置交换机端口类型并允许对应的策略VLAN通过。

```
[SwitchA] interface gigabitethernet 1/0/1
[SwitchA -GigabitEthernet1/0/1] port link-type hybrid
```

[SwitchA -GigabitEthernet1/0/1] port hybrid Untagged vlan 2

[SwitchA -GigabitEthernet1/0/1] quit

[SwitchA] interface gigabitethernet 1/0/2

[SwitchA -GigabitEthernet1/0/2] port link-type trunk

[SwitchA -GigabitEthernet1/0/2] port trunk allow-pass vlan 3

[SwitchA -GigabitEthernet1/0/2] quit

SwitchB上的配置：

SwitchB上的配置与SwitchA上的配置基本类似，具体如下。

<HUAWEI>system-view

<HUAWEI>sysname SwitchB

[SwitchB] vlan batch 2 3

[SwitchB] vlan 2

[SwitchB -vlan2] policy-vlan mac-address 2-2-2 ip 2.2.2.2 gigabitEthernet1/0/1priority 7

[SwitchB -vlan2] quit

[SwitchB] vlan 3

[SwitchB -vlan20] policy-vlan mac-address 4-4-4 ip 4.4.4.4 gigabitEthernet1/0/2priority 5

[SwitchB-vlan20] quit

[SwitchB] interface gigabitethernet 1/0/1

[SwitchB -GigabitEthernet1/0/1] port link-type hybrid

[SwitchB -GigabitEthernet1/0/1] port hybrid Untagged vlan 2

[SwitchB -GigabitEthernet1/0/1] quit

[SwitchB] interface gigabitethernet 1/0/2

[SwitchB -GigabitEthernet1/0/2] port link-type trunk

[SwitchB -GigabitEthernet1/0/2] port trunk allow-pass vlan 3

[SwitchB -GigabitEthernet1/0/2] quit

通过以上配置MAC地址为1-1-1，IP地址为1.1.1.1的用户被自动划分到VLAN 2中，并且只能接在SwitchA上的GE1/0/1端口上；MAC地址为2-2-2，IP地址为2.2.2.2的用户也被自动划分到VLAN 2中，并且只能接在SwitchB上的GE1/0/1端口上，否则将退出VLAN 2。而MAC地址为3-3-3，IP地址为3.3.3.3的用户被自动划分到VLAN 3中，并且只能接在SwitchA上的GE1/0/2端口上；MAC地址为4-4-4，IP地址为4.4.4.4的用户也被自动划分到VLAN 3中，并且只能接在SwitchB上的GE1/0/2端口上，否则将退出VLAN 3。

6.7 VLAN配置管理和典型故障分析与排除

本节首先要介绍一些常见的 VLAN 管理命令，包括一系列用来查看 VLAN 信息的display命令，清除VLAN统计信息的reset命令。然后介绍两个常见的VLAN方面的故 障分析与排除方法。

6.7.1 常见VLAN管理命令

在完成了以上介绍的各种VLAN划分后，可通过在任意视图下执行以下display命令查看对应的VLAN配置信息，验证配置是否成功。还可通过以下reset用户视图命令清除指定VLAN中的统计信息。

（1）**display vlan**： 查看所有VLAN或指定VLAN的显示信息。

（2）**display mac-vlan** {**mac-address** {**all** |mac-address } | **vlan** vlan-id }：查看基于MAC地址划分VLAN的相关信息。

（3）**display ip-subnet-vlan vlan** {**all** | vlan-id1 [**to** vlan-id2] }：查看VLAN上所配置的IP子网信息。

（4）**display protocol-vlan vlan** { **all** | vlan-id1 [**to** vlan-id2] }：查看VLAN上所配置的协议及协议索引信息。

（5）**display protocol-vlan interface** {**all** | interface-type interface-number }：查看接口关联基于协议VLAN划分的配置信息。

（6）**display policy-vlan** { **all | vlan** vlan-id }：查看策略VLAN的配置信息。

（7）**reset vlan** vlan-id **statistics**：清除指定VLAN的报文统计信息。

6.7.2 典型故障分析与排除

在VLAN的配置与使用中，经常会遇到以下两种VLAN方面的故障，下面分别介绍它们的故障原因及排除方法。

1．VLAN内主机不能互通

我们知道，VLAN是用来隔离用户的二层通信的，在不同的VLAN中不能直接通信，但在同一VLAN内部的用户主机是可以直接通信的。但是有时会出现即使是同一VLAN内的用户主机也不能直接通信。

我们先分析一下造成同一VLAN内用户主机不能互通的原因有哪些。网络通信只涉及OSI/RM体系结构中的最低三层，我们一层层来分析。

（1）物理层是一切网络通信的基础，在这层最有可能导致不能通信的原因就是用户所连接的交换机端口没有启用，造成通信线路不通。

（2）再从数据链路层来分析。同一VLAN是直接通过数据链路层的MAC地址进行寻址的，如果交换机错误地学习了某用户的 MAC 地址，则可能造成不能正确通信。或者用户主机上配置了错误的ARP静态表项，导致对应用户主机不能正确地与网络连接。

（3）再从网络层来分析。尽管在VLAN内部是通过数据链路层MAC地址进行寻址的，但来自网络应用的用户数据在经过网络层时封装了源和目的主机的IP地址，如果源和目的用户主机的IP地址不在同一网段，则数据包在到达网络层后直接发到网关，如果网关不通则两用户自然不能彼此通信了。

（4）是否配置了这些用户的隔离（即MUX VLAN功能）。

根据以上分析，可以采取由简到繁的步骤依次排除同一VLAN内用户不能互通的故障。

（1）检查VLAN内需要互通的用户所连交换机端口的状态是否为Up。可在任意视 图下执行display interface interface-type interface-number命令来查看。如果接口的状态为Down，请先排除接口Down的故障；如果成员口的状态是Up，则继续下面的步骤。

（2）检查需要互通的用户主机的IP地址是否在同一网段，如果不是请修改为同一网段，如果故障仍然存在，则继续下面的步骤。

（3）检查设备上MAC地址表项是否正确。在交换机上执行display mac-address命令检查设备学习到的MAC地址、MAC地址对应接口、所属VLAN是否正确，如果不正确请在系统视图下执行 undo mac-addressmac-addressvlan vlan-id命令删除错误的MAC地址表项，并使交换机重新学习指定的MAC地址。

（4）执行完上述操作后，再检查设备学习到MAC地址、MAC地址对应接口、所属VLAN是否正确：如果不正确请继续执行本步检查VLAN配置是否正确，如对应的VLAN是否创建，用户端口是否正确加入了同一个VLAN中等，可通过display vlan vlan-id命令来查看。

（5）如果通过以上排查，用户仍无法互相访问，则检查设备上是否配置了端口隔离。可在系统视图下

执行 interface interface-type interface-number进入故障接口视图，然后执行display this命令查看接口是否配置了端口隔离。如果配置了端口隔离，可使用undo port-isolate enable命令取消端口隔离配置。

（6）取消端口隔离后如果故障依然存在，则检查终端设备上是否配置了错误的静态ARP表项，如果终端设备上配置了错误的静态ARP表项请修正。

2．VLANIF接口Down

另一个典型VLAN故障现象就是VLANIF接口处于Down（关闭）状态。但这种故障比较好排除，主要原因及排除方法如下。

（1）没有交换机端口加入该VLAN中：将对应的交换机端口加入该VLAN中。

（2）加入该VLAN的各交换机端口的物理状态全是Down：排除加入的交换机端口Down状态的原因。一个VLAN中，只要有一个交换机端口的物理状态是Up状态，则该VLANIF接口的状态就是Up状态。

（3）VLANIF接口下没有配置IP地址：VLANIF接口是三层逻辑接口，必须要配置IP地址才能激活的。在该VLANIF接口视图下通过 ip address命令为该VLANIF配置 IP地址。

（4）VLANIF接口被手动关闭：在该VLANIF接口视图下，执行undo shutdown命令开启当前VLANIF接口。

6.8 GVRP配置与管理

本章前面介绍的各种VLAN划分方法都是基于手动来创建各交换机上的各个VLAN的，这对于网络中交换机数量不多，所需划分的VLAN数量也不多时没什么问题，但如果网络交换机数量较大，需要划分的VLAN 也比较多时，管理员需承担较大的 VLAN创建工作量，而且这些工作都是低技术含量的简单重复劳动。为此，VLAN技术的开发者就想到了一种自动注册方式，对应GVRP（GARP VLAN Registration Protocol，GARP VLAN 注册协议）协议。通过它只需在其中一个交换机中创建所需的 VLAN，然后通过自动注册功能在网络中其他交换机中自动创建所需的 VLAN，大大减轻了网络管理员的工作量，也减少了错误出现的机率。

图6-12中所有设备都使能GVRP功能，设备之 间相连的端口均为Trunk端口，并允许所有VLAN通过。通过GVRP只需在SwitchA和SwitchC上分别手工配置静态VLAN100～VLAN1000，设备SwitchB就可以学习到这些VLAN，最后各设备上都存在VLAN100～1000。

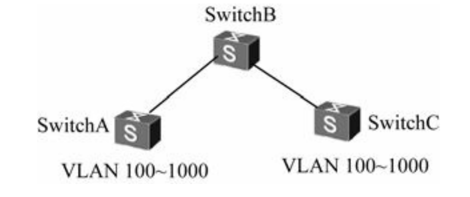


图6-12 GVRP应用示例

本功能除了S1700系列交换机外，其他华为S系列交换机都支持。

6.8.1 GVRP基础

GVRP是GARP（Generic Attribute Registration Protocol，通用属性注册协议）的一种应用，是一种我们通常所说的VLAN中继。通过它，在一个交换机上的VLAN配置可以自动在网络中其他交换机上自动注册，可大大减轻在不同交换机上重复创建VLAN的工作量。如我们在 SwitchA上创建了VLAN 2～10这 9个

VLAN，通过GVRP就可以在网络中其他交换机中自动创建这9个VLAN，相应VLAN会自动添加到 Trunk 端口的允许列表中，当然也可以选择性地自动注册或加入个别 VLAN配置。

GVRP基于GARP机制，主要用于维护设备动态VLAN信息。通过GVRP协议，一台设备上的VLAN信息会迅速传播到整个交换网。GVRP实现动态分发、注册和传播VLAN信息，从而达到减少网络管理员的手工配置量及保证VLAN配置正确的目的。但**GVRP**注册功能仅可在连接网络设备的**Trunk**端口上使能，所以用户计算机所连接的端口仍不能通过**GVRP**功能自动加入到所需的**VLAN**中，仍需要采取手动配置 。这一点要特别注意。

在交换机设备上，每一个参与协议的端口可以视为一个应用实体。当GVRP在设备上启动的时候，每个启动GVRP的端口对应一个GVRP应用实体。

1．VLAN注册和VLNA注销

GVRP协议可以实现VLAN信息的自动注册（Register）和注销（Deregister）：VLAN的注册指的是将端口加入VLAN；VLAN的注销指的是将端口退出VLAN。GVRP协议通过声明和回收声明实现 VLAN 信息的注册和注销。当端口接收到一个 VLAN 信息声明时，该端口将注册该声明中包含的VLAN信息（端口加入VLAN）；当端口接收到一个 VLAN 信息的回收声明时，该端口将注销该声明中包含的 VLAN 信息（端口退出VLAN）。但要注意的是，GVRP协议的属性注册和注销仅仅是对于接收到GVRP协议报文的端口而言的。

2．GARP消息类型

在GARP应用实体之间的信息交互过程中主要有三类消息起作用，分别为Join消息、Leave消息和LeaveAll消息。

（1）Join（加入）消息。当一个GVRP应用实体希望其他设备注册自己的属性信息 时，对外发送Join消息；当收到其他实体发来的Join消息，或本设备静态配置了某些属性，需要其他GARP应用实体进行注册时，也会向外发送Join消息。

Join消息又分为JoinEmpty和JoinIn两种。JoinEmpty消息用来对外声明一个发送该声明者自己还没有注册的属性；而JoinIn消息用来对外声明一个发送该声明者自己已经注册的属性。

（2）Leave（注销）消息。当一个 GARP 应用实体希望其他设备注销自己的某个属性 时，它将对外发送Leave消息；当收到其他实体的Leave消息注销某些属性，或静态注销了某些属性后，也会向外发送Leave消息。

Leave消息也分为LeaveEmpty和LeaveIn两种。LeaveEmpty消息用来注销一个发送该消息者自己没有注册的属性；LeaveIn 消息用来注销一个发送该消息者自己已经注册的属性。

（3）LeaveAll（全部注销）消息。每个应用实体启动后，将同时启动LeaveAll定时器，当该定时器超时后应用实体将对外发送LeaveAll消息。LeaveAll消息用来注销所有属性 ，以使其他应用实体重新注册本实体上所有的属性信息，以此来周期性地清除网络中的垃圾属性（例如某个属性已经被删除，但由于设备突然断电，并没有发送Leave消息来通知其他实体注销此属性）。

3．GARP定时器

GARP协议中用到了 Join、Hold、Leave和LeaveAll 4个定时器。

（1）Join（加入）定时器。Join 定时器是用来确保 Join 消息（包括 JoinIn 消息和JoinEmpty消息）可靠发送。

为了保证一个GARP应用实体发送的Join消息能够可靠地传输到其他应用实体，在发送第一个Join消息后将启动一个Join定时器，如果在一个Join定时器时间内收到了返回的JoinIn消息（表明已成功注册某属性），则不发送第二个Join消息；如果没收到，则再发送一个Join消息。每个GVRP端口维护独立的Join定时

器。

（2）Hold（保持）定时器。Hold定时器是用来控制 Join消息（包括 JoinIn消息和JoinEmpty消息）和Leave消息（包括LeaveIn消息和LeaveEmpty消息）的发送的。

当在GARP应用实体上配置属性或应用实体接收到消息时不会立刻将该消息传播到其他设备，而是在等待一个Hold定时器后再发送消息，设备将此Hold定时器时间段内接收到的Join消息或Leave消息尽可能封装成最少数量的报文，这样可以减少报文的发送量。如果没有Hold定时器的话，每来一个消息就发送一个，造成网络上报文量太大，既不利于网络的稳定，也不利于充分利用每个报文的数据容量。

每个端口维护独立的Hold定时器，但Hold定时器的值要小于等于Join定时器值的一半。

（3）Leave（注销）定时器。Leave定时器是用来控制属性注销的。每个应用实体接收到来自其他的一个应用实体的Leave消息或LeaveAll消息后会启动Leave定时器，如果在Leave定时器超时之前没有接收到该属性的Join消息（可以是来自其他任何应用实体的），属性才会被注销。这是因为网络中如果有一个实体因为不存在某个属性而发送了Leave 消息，并不代表所有的实体都不存在该属性了，因此不能立刻注销属性，而是要 等待其他实体的消息。

例如，某个属性在网络中有两个源，分别在应用实体A和B上，其他应用实体通过协议注册了该属性。当把此属性从应用实体A上删除的时候，实体A发送Leave消息，由于实体B上还存在该属性源，在接收到Leave消息之后，会发送Join消息，以表示它还有该属性。其他应用实体如果收到了应用实体B发送的Join消息，则该属性仍然被保留，不会被注销。只有当其他应用实体等待两个Join定时器以上仍没有收到该属性的Join消息时，才认为网络中确实没有该属性了。

每个端口维护独立的Leave定时器，但要求Leave定时器的值大于2倍Join定时器的值。

（4）LeaveAll（全部注销）定时器。每个GARP应用实体启动后，将同时启动LeaveAll定时器，当该定时器超时后 GARP 应用实体将对外发送 LeaveAll 消息，随后再启动LeaveAll定时器，开始新的一轮循环。

接收到LeaveAll消息的实体将重新启动所有的定时器，包括LeaveAll定时器。在自己的LeaveAll定时器重新超时之后才会再次发送LeaveAll消息，这样就避免了短时间内发送多个LeaveAll消息。

如果不同设备的LeaveAll定时器同时超时，就会同时发送多个LeaveAll消息，增加不必要的报文数量。为了避免这种情况发生，实际定时器运行的值大于LeaveAll定时器的值，小于1.5倍LeaveAll定时器值的一个随机值。一次LeaveAll事件相当于对全网所有属性的一次Leave（注销）。由于LeaveAll影响范围很广，所以建议LeaveAll定时器的值不能太小，至少应该大于Leave定时器的值。

每个设备只在全局维护一个LeaveAll定时器。

4．注册模式

我们可以把手工配置的VLAN称为静态VLAN，通过GVRP协议创建的VLAN称为动态VLAN。GVRP有以下3种注册模式，它们对静态VLAN和动态VLAN的处理方式各不相同。

（1）Normal模式：允许该端口动态注册、注销VLAN，传播动态VLAN和静态VLAN信息。这是最常用的一种动态注册模式，也是唯一一种真正具有动态注册VLAN功能的模式。

（2）Fixed模式：禁止该端口动态注册、注销VLAN，只传播静态VLAN信息，不传播动态VLAN信息。也就是说被设置为Fixed模式的Trunk端口，即使允许所有VLAN通过，实际通过的VLAN也只能是手动创建的那部分。

（3）Forbidden模式：禁止该端口动态注册、注销VLAN，不传播除VLAN1以外的任何的VLAN信息。也就是说被配置为Forbidden模式的Trunk端口，即使允许所有VLAN通过，实际通过的VLAN也只能是VLAN1。

6.8.2 GVRP工作原理

下面通过一个简单的例子来介绍一下GVRP的工作过程。该例子分4个阶段描述了一个VLAN信息在网络中是如何被注册和注销的。

1．VLAN信息的单向注册

GVRP的VLAN注册是通过Join消息来实现的，一个VLAN信息的成功注册同时需要JoinEmpty和JoinIn这两种消息，**JoinEmpty**相当于注册请求消息，而JoinIn相当于注册成功应答消息。这里所说的单向注册是仅通过 JoinEmpty 消息由发送者到达 网络中其他所有 GARP 应用实体的传递过程来完成的。

下面以图 6-13 所示的结构为例介绍 VLAN信息的单向注册流程。假设在SwitchA上创建了静态VLAN2，现要通过GVRP的VLAN信息单向注册功能，将SwitchB和SwitchC的相应端口自动加入VLAN2。此时GVRP的VLAN信息的单向注册流程如下。

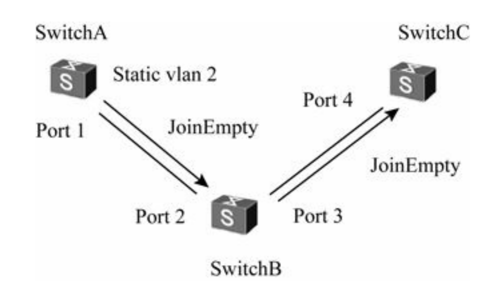


图6-13 VLAN信息的单向注册示例

（1）在SwitchA上创建静态VLAN 2后，因为已发生了VLAN信息变化，所以使能了 GVRP 功能的 Port1 会启动 Join 定时器和 Hold 定时器。等待 Hold 定时器超时后， SwitchA向SwitchB发送第一个JoinEmpty消息（虽然此时Port1已加入VLAN 2中，但因为在SwitchA上VLAN 2是静态创建的，而不是动态注册的，所以仍以 JoinEmpty消息发送）。Join定时器超时后再次启动Hold定时器，再等待Hold定时器超时后，向SwitchB发送第二个JoinEmpty消息。

（2）SwitchB在收到第一个来自 SwitchA的 JoinEmpty消息后创建动态VLAN 2，并把接收到 JoinEmpty消息的 Port2加入动态VLAN 2中。同时告知其 Port3启动 Join定时器和Hold定时器，等待Hold定时器超时后向SwitchC发送第一个JoinEmpty消息（因为此时 Port3也还没加入VLAN 2中）。同样在 Join定时器超时后再次启动Hold定时器，Hold定时器超时之后，向SwitchC发送第二个JoinEmpty消息。SwitchB上收到来自SwitchA的第二个 JoinEmpty时，因为此时Port2已经加入动态VLAN 2，所以不作处理。

（3）SwitchC在收到来自SwitchB的第一个 JoinEmpty消息后也创建动态VLAN 2，并把接收到JoinEmpty消息的Port4加入动态VLAN 2中。同样，当SwitchC收到来自SwitchB的第二个 JoinEmpty后，因为Port4已经加入动态VLAN 2，所以也不作处理。

此后，每当 Leaveall 定时器超时或收到 LeaveAll 消息时，设备会重新启动Leaveall 定时器、Join 定时器、Hold 定时器和 Leave 定时器。SwitchA 的 Port1 在Hold定时器超时之后发送第一个JoinEmpty消息，再等待Join定时器+Hold定时器之后，发送第二个JoinEmpty消息，SwitchB向SwitchC发送JoinEmpty消息的过程也是如此。

以上就是VLAN信息的单向注册过程，是由JoinEmpty消息单向（注意消息发送的方向）传递的过程，但还没有完成整个VLAN 2属性的注册，还需要进行下面将要介绍的VLAN信息双向注册过程。

2．VLAN信息的双向注册

通过上述VLAN信息的单向注册过程，Port1、Port2、Port4已经加入VLAN2，但是 Port3还没有加入**VLAN** 2（只有收到 **JoinEmpty**消息或 **JoinIn**消息的端口才能 加入动态**VLAN**，而Port3并没收到这些消息）。为使 VLAN 2流量可以双向互通，还需要进行SwitchC到SwitchA方向的VLAN信息的注册过程。具体流程如下（见图6-14）。

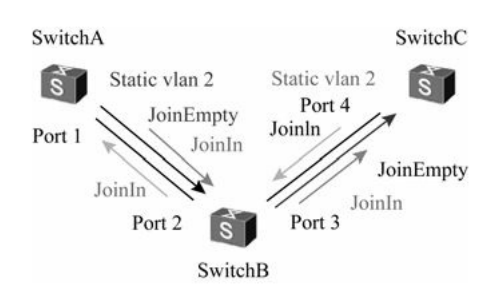


图6-14 VLAN信息的双向注册

（1）VLAN信息的单向注册完成后，在SwitchC上创建静态VLAN 2，此时会将动态VLAN转换成静态VLAN。Port4启动Join定时器和Hold定时器，等待Hold 定时器超时后，SwitchC 向 SwitchB 发送第一个JoinIn消息（因为 Port4已经注册了VLAN 2，所以发送 JoinIn消息）， Join 定时器超时后再次启动 Hold 定时器，Hold 定时器超时之后，向 SwitchB发送第二个JoinIn消息。

（2）SwitchB在收到来自SwitchC第一个JoinIn消息后，把接收到JoinIn消息的Port3加入动态 VLAN 2中。同时告知 Port2启动 Join定时器和Hold定时器，等待Hold 定时器超时后，向 SwitchA 发送第一个 JoinIn消息。Join 定时器超时后再次启动Hold定时器，Hold定时器超时之后，向SwitchA发送第二个JoinIn消息。SwitchB收到来自 SwitchC的第二个 JoinIn消息后，因为 Port3已经加入动态VLAN 2，所以不作处理。

（3）SwitchA在收到来自SwitchB的JoinIn消息之后，停止向SwitchB发送JoinEmpty消息。此后，当Leaveall定时器超时或收到LeaveAll消息，设备重新启动Leaveall定时器、Join定时器、Hold定时器和Leave定时器。

（4）SwitchA的Port1在Hold定时器超时之后就开始向SwitchB发送JoinIn消息。SwitchB的Port3也会向SwitchC发送JoinIn消息。

（5）SwitchC在收到来自SwitchB的JoinIn消息后，由于本身已经创建了静态VLAN2，所以不会再创建动态VLAN2。

从以上流程可以看出，双向注册过程中发送的是JoinIn消息，而且是一个双向、封闭环路过程（注意图中JoinIn消息的发送方向）。

3．VLAN信息的单向注销

VLAN信息的消息过程是使用Leave消息或者LeaveEmpty和LeaveIn消息来实现的。

同样以上面的示例进行介绍，当设备上不再需要VLAN2时，可以通过VLAN信息的注销过程将VLAN2从设备上删除。下面以图 6-15为例介绍SwitchA上删除VLAN 2 后在其他路由器上的单向注销过程。

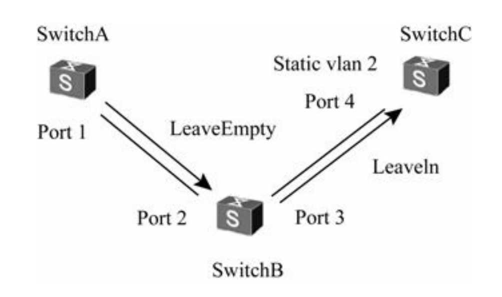


图6-15 VLAN信息的单向注销

（1）在SwitchA上删除静态VLAN 2，因为VLAN信息发生了变化，Port1启动Hold定时器，等待Hold定时器超时后，SwitchA向SwitchB发送 LeaveEmpty 消息（同样是因为 Port1 不是动态加入VLAN 2的）。**LeaveEmpty**消息只需发送一次 。

（2） SwitchB在收到来自SwitchA的LeaveEmpty 消息后，Port2启动Leave定时器，等待Leave定时器超时之后 Port2注销VLAN 2，将 Port2从动态VLAN2中删除（由于此时VLAN 2中还存在 Port3，所以不能直接删除VLAN 2）。同时告知 Port3 启动Hold定时器和 Leave定时器，等待Hold定时器超时后，向 SwitchC发送LeaveIn消息（因为 Port3是动态加入VLAN 2的）。由于SwitchC的静态VLAN 2还没有删除，Port3在 Leave定时器超时之前仍然能够收到Port4 发送的 JoinIn 消息，所以此时 SwitchA 和 SwitchB 上仍然能够学习到动态的VLAN 2。

（3）SwitchC在收到来自SwitchB的LeaveIn消息后，由于SwitchC上存在静态VLAN 2 （此时VLAN 2已转换成静态的），所以Port4也不会从VLAN 2中删除。

通过以上单向注销过程可以发现，只有Port1、Port2注销了VLAN 2，Port3和Port4都还没有注销VLAN 2。

4．VLAN信息的双向注销

为了彻底删除所有设备上的VLAN2，需要进行VLAN信息的双向注销。下面是双向注销的流程（见图6-16）。

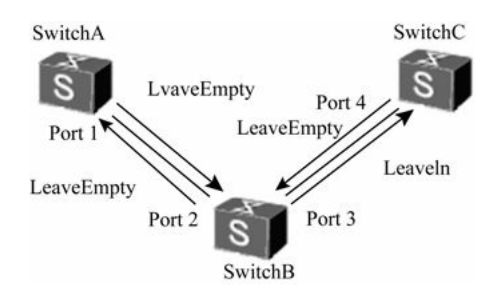


图6-16 VLAN信息的双向注销

（1）在 SwitchC上手动删除静态VLAN 2， Port4启动Hold定时器，等待Hold定时器超时后，SwitchC会向SwitchB发送LeaveEmpty消息。

（2）SwitchB在收到来自SwitchC的LeaveEmpty消息后，Port3启动Leave定时器，等待Leave定时器超时之后Port3注销VLAN 2，将Port3从动态VLAN 2中删除并删除动态VLAN 2，同时告知Port2启动Hold定时器，等待Hold定时器超时后，向SwitchA发送LeaveEmpty消息。

（3）SwitchA在收到来自SwitchB的LeaveEmpty消息后，Port1启动Leave定时器，等待Leave定时器超时之后Port1注销后面学习到的动态VLAN 2，将Port1从动态VLAN 2中删除并删除动态VLAN 2。

通过以上过程就完成了整个VLAN 2的注销过程。

6.8.3 使能GVRP功能

支持GVRP的设备都有如表6-8所示的缺省参数配置，但可以修改这些缺省配置。

表6-8 华为交换机的缺省GVRP参数配置

参数	缺省值
GVRP 功能	全局和接口的 GVRP 功能都处于关闭状态
GVRP 接口注册模式	normal
LeaveAll 定时器	1000 厘秒
Hold 定时器	10 厘秒
Join 定时器	20 厘秒
Leave 定时器	60 厘秒

本节先介绍使能GVRP功能的配置方法，其他参数的配置将在后面小节介绍。

GVRP 功能的使能有两个层次，一个是整个交换机全局使能，另一个是在具体的交换机端口上使能。但在使能端口的GVRP功能之前，必须先全局使能GVRP功能。另外， **GVRP**功能只能配置在**Trunk**类型的接口上 ，并且需要保证所有需要动态注册的VLAN都能够从该端口通过。具体配置方法如表6-9所示。

表6-9 使能GVRP功能的配置步骤

步骤	命令	说明
1	**system-view** 例如：<HUAWEI> system-view	进入系统视图
2	**gvrp** 例如：[HUAWEI] **gvrp**	全局使能 GVRP 功能。缺省情况下，全局和接口的 GVRP 功能都处于关闭状态，可用 **undo gvrp** 命令全局关闭 GVRP 功能
3	**interface** *interface-type interface-number* 例如：[HUAWEI] **interface** ethernet **0/0/1**	键入要启用 GVRP 功能的交换机端口
4	**port link-type trunk** 例如：[HUAWEI-Ethernet0/0/1] **port link-type trunk**	设置以上交换机端口为 Trunk 类型
5	**port trunk allow-pass vlan** { { *vlan-id1* [**to** *vlan-id2*] } &<1-10> | **all** } 例如：[HUAWEI-Ethernet0/0/1] **port trunk allow-pass vlan** 2 **to** 10	配置允许在其他交换机上动态注册的 VLAN 通过。其他说明参见表 6-3 第 7 步 当设备GARP定时器使用缺省值时最多支持256个动态VLAN，使用推荐值时最多支持 4 094 个动态 VLAN
6	**gvrp** 例如：[HUAWEI-Ethernet0/0/1]**gvrp**	在以上交换机端口上使能 GVRP 功能。缺省情况下，全局和接口的 GVRP 功能都处于关闭状态，可使用 **undo gvrp** 命令关闭接口上的 GVRP 功能

6.8.4 配置GVRP端口注册模式

在启用了GVRP功能的交换机端口上，可以配置**normal**、**fixed** 和**forbidden** 3种通过GVRP在其他交换机上动态注册VLAN的注册模式（参见6.8.1节，但这是可选配置任务 ）：

在启用了GVRP功能的Trunk端口上配置注册模式的方法很简单，只需要在对应的接口视图下通过 **gvrp registration** { **fixed** | **forbidden** |**normal** }命令配置即可。命令中的3个多选一选项分别对应以上3种端口注册模式。缺省情况下，GVRP接口注册模式为**normal** 模式，可用**undo gvrp registration** 命令恢复GVRP接口注册模式为缺省的**normal** 模式。

配置端口注册模式前需要全局和端口均使能**GVRP**功能，且配置端口类型为**Trunk**类型 。

【示例 】设置Gigabitethernet 0/0/1 GVRP端口注册模式为Fixed模式。

```
<HUAWEI>system-view
[HUAWEI] interface gigabitethernet 0/0/1
[HUAWEI-GigabitEthernet0/0/1] port link-type trunk
[HUAWEI-GigabitEthernet0/0/1] gvrp
[HUAWEI-GigabitEthernet0/0/1] gvrp registration fixed
```

6.8.5 配置GARP定时器参数值

在一台交换机设备使能了GARP注册功能后，将同时启动LeaveAll定时器，当该定时器超时后，该交换机将对外发送LeaveAll消息，以使其他使能了GARP功能的交换机重新注册本交换机上所有的属性信息。随后再启动LeaveAll定时器，开始新的一轮循环。

【经验之谈】在网络中有多台交换机的情况下，各个交换机的 LeaveAll 定时器的取值可能不相同，此时每台交换机都将以整个网络中配置的最小LeaveAll定时器值来发送LeaveAll消息。因为每次LeaveAll定时器超时后都会发送LeaveAll消息，其他的交换机在接收到这个LeaveAll消息后都会清零LeaveAll定时器。所以即使整个网络中存在很多不同的LeaveAll定时器，实际上也只有最小的那个LeaveAll定时器起作用。

除了可以配置LeaveAll定时器参数外，还可以配置在6.8.1节介绍的其他定时器参数，如Join定时器、Hold定时器和Leave定时器。但要注意的是，各个定时器的取值范围会由于其他定时器取值的改变而改变。如果用户想要设置的定时器的值不在当前可以设置的取值范围内，可以通过改变相关定时器的取值实现。如果用户想恢复各定时器的值为缺省值，可以先恢复 Hold 定时器的值为缺省值，然后再依次恢复 Join、Leave、LeaveAll定时器的值为缺省值。当然这也是可选配置任务，因为这些定时器参数都有它们的缺省值。

在实际组网中，建议用户将GVRP定时器配置为以下的推荐值。

（1）Hold定时器：100厘秒（1秒钟）。

（2）Join定时器：600厘秒（6秒钟）。

（3）Leave定时器：3000厘秒（30秒钟）。

（4）LeaveAll定时器：12000厘秒（2分钟）。

当动态VLAN超过100个或运行GVRP的网络超过3台设备时，需将定时器配置为推荐值。当动态VLAN数或设备数增加时，定时器的时间也需要相应地增加。

以上各GARP定时器参数的配置步骤如表6-10所示。

表6-10 GARP定时器参数的配置步骤

步骤	命令	说明
1	**system-view** 例如：<HUAWEI> system-view	进入系统视图
2	**garp timer leaveall** *timer-value* 例如：[HUAWEI] **garp timer leaveall** 2000	全局配置 GARP 的 LeaveAll 定时器的值。参数的 *timer-value* 取值范围为 65～32 765 的整数，单位为厘秒，取值必须是 5 厘秒的倍数，且 LeaveAll 定时器的值应大于所有端口 Leave 定时器的值 缺省情况下，LeaveAll 定时器的值为 1000 厘秒，即 10 秒。可用 **undo garp timer leaveall** 命令恢复 GARP 的 LeaveAll 定时器为缺省值 由于各交换机端口 Leave 定时器的值受全局 LeaveAll 定时器的值限制，所以在配置 LeaveAll 定时器的值时，需要保证设备上所有配置 GARP 定时器的端口都处于正常工作状态

（续表）

步骤	命令	说明
3	**interface** *interface-type interface-number* 例如：[HUAWEI] **interface** ethernet **0/0/1**	键入要启用 GVRP 功能的交换机端口
4	**garp timer** { **hold** | **join** | **leave** } *timer-value* 例如：[HUAWEI-Ethernet0/0/1] **garp timer hold** 200	配置交换机端口的 **Hold** 定时器、**Join** 定时器、**Leave** 定时器值 （1）当配置 **Hold** 定时器值时，参数 *timer-value* 的取值下限为 10 厘秒；取值上限小于等于 1/2 Join 定时器的值，可以通过改变 Join 定时器的取值改变；取值必须是 5 厘秒的倍数 （2）当配置 **Join** 定时器值时，参数 *timer-value* 的取值下限为大于等于 2 倍 Hold 定时器的值，可以通过改变 Hold 定时器的取值实现；取值上限为小于 1/2 Leave 定时器的取值，可以通过改变 Leave 定时器的取值改变；但取值也必须是 5 厘秒的倍数 （3）当配置 **Leave** 定时器值时，参数 *timer-value* 的取值下限为大于 2 倍 Join 定时器的值，可以通过改变 Join 定时器的取值改变；取值上限为小于 LeaveAll 定时器的值，可以通过改变 LeaveAll 定时器的取值改变；取值也必须是 5 厘秒的倍数 缺省情况下，**Hold** 定时器的值为 10 厘秒，**Join** 定时器的值为 20 厘秒，**Leave** 定时器的值为 60 厘秒，可用 **undo garp timer** { **hold** | **join** | **leave** } [*timer-value*]命令恢复对应交换机端口的对应 GARP 定时器的值为缺省值

6.8.6 GVRP配置管理

配置好GVRP功能后，可在任意视图下使用以下相关display命令进行配置管理或相关信息查看，在用户视图下使用以下reset命令清除GVRP相关统计信息。

（1）使用**display gvrp status** 命令查看全局GVRP功能的使能或去使能状态信息。

（2）使用**display gvrp statistics** [**interface** { interface-type interface-number [to interface- type interface-number] }&<1-10>] 命令查看特定交换机端口上的GVRP统计信息。

（3）使用**display garp timer** [**interface** { interface-type interface-number [to interface- type interface-number] }&<1-10>] 命令查看特定交换机端口上配置的GARP定时器值。

（4）使用 **reset garp statistics** [**interface** { interface-type interface-number [to interface- type interface-number] }&<1-10>] 命令清除特定交换机端口上的GARP统计信息。

6.8.7 GVRP配置示例

本示例拓扑结构如图 6-17所示，公司A（Company A）、公司A的分公司（Branch of Company A）以及公司B（Company B）之间有较多的交换设备相连，需要通过GVRP功能实现VLAN的动态注册。公司A的分公司通过SwitchA和SwitchB与公司A互通；公司B通过SwitchB和SwitchC与公司A互通，但只允许公司B配置的VLAN通过。

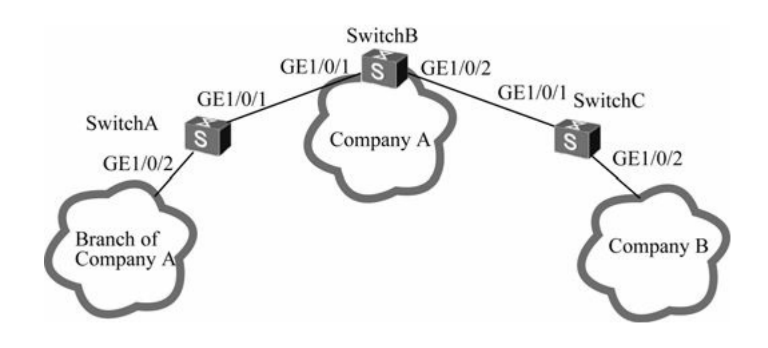


图6-17 GVRP配置示例拓扑结构

1．配置思路分析

本示例中有两个明确的要求，那就是公司A的分公司与公司A之间要求互联互通，所以在VLAN动态注册上没有限制；而公司B与公司A之间的连接仅允许公司B上静态配置的VLAN通过。针对以上两方面的要求，可以采用如下的思路来配置各交换机上的GVRP功能。

（1）在公司A、公司A的分公司和公司B网络中的各交换机Trunk端口上使能GVRP功能，并配置这些端口的注册模式为Normal，实现VLAN的动态注册。

（2）在SwitchC上手动创建整个网络中所需的静态VLAN（假设为VLAN 101～200）。

（3）在SwitchA与公司A的分公司、SwitchB连接的Trunk端口，以及SwitchC与公司B连接的Trunk端口上配置GVRP功能，并配置这些端口的注册模式为Normal。

说明

通过以上三项配置任务就可以使得公司 A、公司 A 的分公司和公司 B，以及SwitchA和SwitchB能动态注册来自SwitchC上配置的静态VLAN。

（4）在SwitchC与SwitchB连接的Trunk端口上配置GVRP功能，并配置注册模式为Fixed，其目的就是要禁止在该端口上动态注册来自公司A网络、公司A的分公司网络，以及SwitchA和SwitchB上创建的VLAN，但仍允许通过该端口向外传播GVRP注册消息，以使公司A网络、公司A的分公司网络，以及SwitchA和SwitchB能动态注册来自 SwitchC 上配置的静态 VLAN，最终实现示例中要求的仅允许公司 B 配置的静态VLAN（其实是在SwitchC上静态创建的）与公司A互访的要求。

2．配置步骤

下面是各交换机的具体配置步骤。

SwitchA交换机的配置：

（1）全局使能GVRP功能。

```
<HUAWEI>system-view
[HUAWEI] sysname SwitchA
[SwitchA] gvrp
```

（2）配置与公司 A 的分公司和 SwitchB 相连的端口均为 Trunk 类型，并允许所有VLAN通过。同时使能GVRP功能，并配置GVRP注册模式为Normal。

```
[SwitchA] interface gigabitethernet 1/0/1
[SwitchA-GigabitEthernet1/0/1] port link-type trunk
[SwitchA-GigabitEthernet1/0/1] port trunk allow-pass vlan all
[SwitchA-GigabitEthernet1/0/1] gvrp
[SwitchA-GigabitEthernet1/0/1] gvrp registration normal
[SwitchA-GigabitEthernet1/0/1] quit
[SwitchA] interfacegigabitethernet 1/0/2
[SwitchA-GigabitEthernet1/0/2] port link-type trunk
[SwitchA-GigabitEthernet1/0/2] port trunk allow-pass vlan all
[SwitchA-GigabitEthernet1/0/2] gvrp
[SwitchA-GigabitEthernet1/0/2] gvrp registration normal
[SwitchA-GigabitEthernet1/0/2] quit
```

SwitchB交换机上的配置：

（1）全局使能GVRP功能。

<HUAWEI>system-view

[HUAWEI] sysname SwitchB

[SwitchB] gvrp

（2）配置与SwitchA和SwitchC相连的端口均为Trunk类型，并允许所有VLAN通过。同时使能GVRP功能，并配置GVRP注册模式为Normal。

[SwitchB] interface gigabitethernet 1/0/1

[SwitchB-GigabitEthernet1/0/1] port link-type trunk

[SwitchB-GigabitEthernet1/0/1] port trunk allow-pass vlan all

[SwitchB-GigabitEthernet1/0/1] gvrp

[SwitchB-GigabitEthernet1/0/1] gvrp registration normal

[SwitchB-GigabitEthernet1/0/1] quit

[SwitchB] interfacegigabitethernet 1/0/2

[SwitchB-GigabitEthernet1/0/2] port link-type trunk

[SwitchB-GigabitEthernet1/0/2] port trunk allow-pass vlan all

[SwitchB-GigabitEthernet1/0/2] gvrp

[SwitchB-GigabitEthernet1/0/2] gvrp registration normal

[SwitchB-GigabitEthernet1/0/2] quit

SwitchC交换机上的配置：

（1）全局使能 GVRP 功能，根据需要手动创建所需的 VLAN，如 VLAN101～VLAN200。最终通过GVRP的VLAN注册功能可使公司A、公司A的分公司和公司B的网络中都有这100个VLAN。

<HUAWEI>system-view

[HUAWEI] sysname SwitchC

[SwitchC] vlan batch 101 to 200

[SwitchC] gvrp

（2）配置与SwitchC和公司B连接的端口均为Trunk类型，并允许所有VLAN通过。

[SwitchC] interface gigabitethernet 1/0/1

[SwitchC-GigabitEthernet1/0/1] port link-type trunk

[SwitchC-GigabitEthernet1/0/1] port trunk allow-pass vlan all

[SwitchC-GigabitEthernet1/0/1] quit

[SwitchC] interface gigabitethernet 1/0/2

[SwitchC-GigabitEthernet1/0/2] port link-type trunk

[SwitchC-GigabitEthernet1/0/2] port trunk allow-pass vlan all

[SwitchC-GigabitEthernet1/0/2] quit

（3）使能与SwitchC和公司B连接端口的GVRP功能，并配置与SwitchB相连端口的注册模式为Fixed模式，以便在与公司A的通信仅允许在SwitchC上静态创建的VLAN 101～200这100个VLAN的帧通过；配置与公司B连接的端口的注册模式为Normal，以便公司B网络也能动态注册在SwitchC上静态创建的VLAN 101～200这100个VLAN。

[SwitchC] interface gigabitethernet 1/0/1

[SwitchC-GigabitEthernet1/0/1] gvrp

[SwitchC-GigabitEthernet1/0/1] gvrp registration fixed

[SwitchC-GigabitEthernet1/0/1] quit

[SwitchC] interface gigabitethernet 1/0/2

[SwitchC-GigabitEthernet1/0/2] gvrp

[SwitchC-GigabitEthernet1/0/2] gvrp registration normal

[SwitchC-GigabitEthernet1/0/2] quit

3．验证配置结果

配置完成后，公司A的分公司、公司A和公司B中的VLAN 101～200中同一VLAN内的用户间都可以直接互访。在 SwitchA上使用 display gvrp statistics命令可查看各Trunk端口上的GVRP统计信息，其中包括GVRP状态、GVRP注册失败次数、上一个GVRP数据单元源MAC地址和接口GVRP注册类型，结果如下，结果显示与上面的配置是一致的，证明配置成功。

<SwitchA>display gvrp statistics

GVRP statistics on port GigabitEthernet1/0/1

GVRP status : Enabled

GVRP registrations failed : 0

GVRP last PDU origin : 0000-0000-0000

GVRP registration type : Normal

GVRP statistics on port GigabitEthernet1/0/2

GVRP status : Enabled

GVRP registrations failed : 0

GVRP last PDU origin : 0000-0000-0000

GVRP registration type : Normal

SwitchB和SwitchC的查看方法与SwitchA类似，不再赘述。

6.9 VLAN间通信配置与管理

我们知道，划分VLAN的目的是为了隔离同一网段中各主机间的直接二层通信，以缩小广播域，减小广播风暴产生的可能性和影响。但是在大多数情况下，不同VLAN中的主机又需要相互通信。为了达到既二层隔离，又能相互通信，华为交换机产品中提供了 3种解决方案，那就是通过配置三层VLANIF接口、三层以太网子接口、VLAN Switch实现VLAN间的通信。本节要对VLAN间通信方式、三种VLAN间通信方案及配置方法进行全面介绍。

6.9.1 两种VLAN间通信方式

VLAN间的三层通信存在两种不同的情形：一是相互通信的不同VLAN同处于一个交换机上，二是相互通信的不同VLAN处于不同的交换机上。下面分别介绍这两种VLAN间通信方式。

1．同一台交换机上的VLAN间通信

这种情形的示意图如图 6-18所示。示例中的VLAN 2、VLAN 3和VLAN 4在同一个三层交换机（也可以是路由器）上。此时要实现这3个VLAN间的三层通信只需在该三层交换机为这3个VLAN各自配置VLANIF接口IP地址（要求在不同IP子网中）即 可，因为在华为交换机中IP路由功能是一直启用的，加上这些VLAN

是直接连接在同一台三层交换机上，相当于直连路由，所以无需其他额外配置就可以实现同一台交换机上不同VLAN间的三层互通。

2．不同交换机上的VLAN间通信

这种情形的示意图如图 6-19所示。示例中的VLAN 2、VLAN 3和VLAN 4不仅在同一台三层交换机（也可以是路由器）上有，而且在不同的三层交换机上也有，这就涉及跨三层设备的VLAN间通信问题了。如果相互通信的不同VLAN位于不同的三层交换机上，不仅要为各VLAN配置VLANIF接口IP地址（不同交换机中的相同VLAN各自可以配置一个同网段的 VLANIF接口IP地址），还要在三层设备上配置到达各个VLANIF 接口所在网段的可达路由（可以是静态路由，也可以是各种动态路由）。

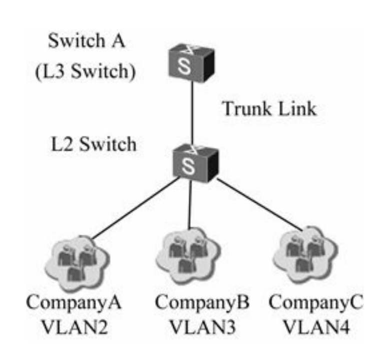


图6-18 同一三层交换机上的VLAN间通信示例

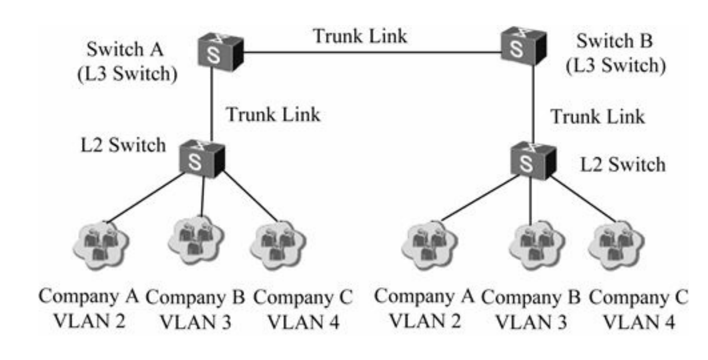


图6-19 跨三层交换机的VLAN间通信示例

6.9.2 VLAN间通信方案及实现原理

在划分VLAN后，不同VLAN之间不能直接进行二层通信。如果要实现VLAN间通信，可以采取以下3种方案之一。

1．三层VLANIF接口方案

这是一种通过计算机网络体系结构中第三层（网络层）来实现VLAN间通信的解决 方案。每个VLAN都可以配置一个三层VLANIF逻辑接口，而这些VLANIF接口就作为对应VLAN内部用户主机的缺省网关，通过三层交换机内部的IP路由功能可以实现同一交换机上不同VLAN的三层互通，不同交换机上不同VLAN间的三层互通需要配置各VLANIF接口所在网段间的路由。

该方案除**S1700**系列外，其他所有华为**S**系列交换机均支持。

在图6-20所示的网络中，Device交换机上划分了两个VLAN：VLAN2和VLAN3。可通过如下配置实现VLAN间互通。

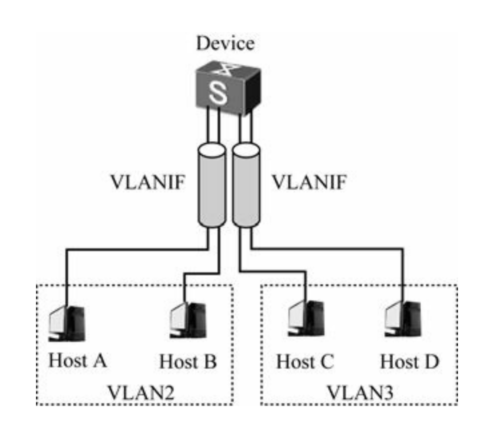


图6-20 通过VLANIF接口实现VLAN间通信的示例

（1）在Device上创建两个VLANIF接口并配置VLANIF接口的IP地址，但这两个VLANIF接口对应的IP地址不能在同一网段。

（2）将各VLAN中的用户设备缺省网关设置为所属VLAN对应VLANIF接口的IP地址。

现在仅以位于VLAN 2中的主机A向位于VLAN 3中的主机C发起通信为例，介绍通过VLANIF接口进行VLAN间三层互通的基本原理。具体通信流程如下。

（1）在主机A向主机C发送的数据包到了网络层后，主机A先将包中的目的IP地址——主机C的IP地址和自己所在网段进行比较。

（2）发现主机C和自己不在同一个子网，于是主机A以广播方式在本子网内发送一个ARP请求帧，其目的是查寻自己的网关——VLANIF2接口的MAC地址。

（3）VLANIF2 接口经过与 ARP 请求帧中的目的 IP 地址进行比较，发现自己的 IP地址与其一致，接收该ARP请求帧，然后以单播方式向主机A返回一个ARP应答帧，帧中的源MAC地址即为VLANIF2的MAC地址。

（4）在主机A接收由VLANIF2接口返回的ARP应答帧后从中学习到了VLANIF2接口的MAC地址。

（5）主机A利用所获得的网关VLANIF2接口的MAC地址，重新进行数据帧封装，把帧中的目的MAC改为VLANIF2接口MAC地址，目的IP仍为主机C的IP地址，然后发送给网关——VLANIF2接口。

（6）Device交换机在收到该数据帧后进行三层转发，发现帧中的目的IP地址——主机C的IP地址为直连路由，数据帧直接通过该主机的网关——VLANIF3接口进行转发。

（7）VLANIF3接口作为VLAN 3内主机的网关，在收到数据帧后如果已有主机 C的 IP地址与MAC地址映射表，则直接发送给主机C，否则VLANIF3接口先在VLAN 3内以广播方式发送一个ARP请求帧，查寻主机C的MAC地址。

（8）主机C在收到ARP广播帧后向VLANIF3接口返回一个ARP应答帧。

（9）VLANIF3接口在收到主机C发来的ARP应答帧后再次进行数据帧封装，把帧中的目的MAC地址改为主机C的真实MAC地址（其他不变），然后把主机A发来的数据帧发送给主机 C。这样主机 A 之后要发给 C 的数据帧都先发送给网关，由网关——VLANIF3接口做三层转发。

主机C与主机A之间的通信原理一样，最终实现VLAN间的三层互通。

2．三层以太网子接口方案

三层以太网子接口是一种同时具备三层以太网物理接口和二层以太网物理接口双重特性的逻辑接口。即它具有三层以太网物理接口的三层路由功能，同时又具有二层以太网物理接口封装 VLAN 标签的特性。通过三层以太网子接口就可以实现不同VLAN间的三层互通，也就是我们通常所说的“单臂路由”，在三层交

换机和路由器中均可实现。

该方案仅 **5700HI**和 **5710EI**子系列、**S7700**、**S9300**和 **S9700**系列华为交换机支持。

如图6-21所示，DeviceA为支持配置子接口的三层设备，DeviceB为二层交换设备。 LAN 通过 DeviceB 的二层以太网接口与DeviceA 的三层以太网接口相连。连接在DeviceB上的用户主机被划分到两个VLAN：VLAN2 和 VLAN3。这时可通过如下配置实现VLAN间互通。

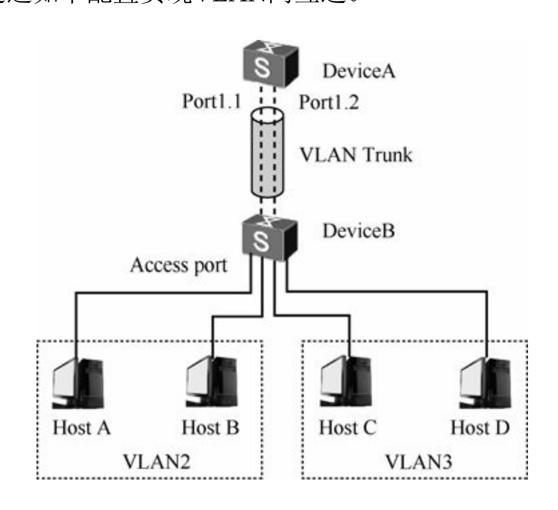


图6-21 通过子接口实现VLAN间通信的示例

（1）在DeviceA与DeviceB相连的三层以太网接口上创建两个子接口 Port1.1 和Port1.2，并配置 802.1Q 封装与 VLAN2 和VLAN3分别对应。

（2）为以上这两个子接口配置与各自所属VLAN对应网段的IP地址。

（3）将 DeviceB 与 DeviceA 相连的二 层以太网接口类型配置为 Trunk或Hybrid类型，并同时允许 VLAN2和 VLAN3的帧通过。

（4）将VLAN 2和VLAN 3中的用户设备的缺省网关设置为所属VLAN对应三层以太网子接口的IP地址。

现在同样以主机A向主机C发起通信为例介绍三层以太网子接口的VLAN间通信方案的基本原理（其实基本过程与前面介绍的VLANIF接口VLAN间通信方案是一样的，只不过这里的网关是各VLAN所对应的子接口）。具体流程如下。

（1）在主机A向主机C发送的数据包到了网络层后，主机A先将包中的目的IP地址——主机C的IP地址和自己所在网段进行比较。

（2）发现主机 C 和自己不在同一个子网，于是主机 A 以广播方式在本子网内发送一个ARP请求帧，其目的是查寻自己的网关VLAN 2对应的Port1.1子接口的MAC地址。

（3）Port1.1子接口经过与ARP请求帧中的目的IP地址进行比较，发现自己的IP地址与其一致，接收该ARP请求帧，然后以单播方式向主机A返回一个ARP应答帧，帧中的源MAC地址即为Port1.1子接口的MAC地址。

（4）主机A接收由Port1.1子接口返回的ARP应答帧后从中学习到该子接口的MAC地址。

（5）主机A利用所获得的网关Port1.1子接口的MAC地址，重新封装数据帧，把目的MAC地址改为Port1.1子接口MAC，目的IP仍为主机C的IP地址，然后发送给网关Port1.1子接口。

（6）DeviceA交换机在收到该数据帧后进行三层转发，发现其目的IP地址——主机C的 IP地址为直连路由，数据帧直接通过该主机的网关——VLAN 3对应Port1.2子接口进行转发。

（7）Port1.2子接口作为VLAN 3内主机的网关，在收到数据帧后如果已有主机C的IP地址与MAC地址映射表，则直接发送给主机C，否则Port1.2子接口先在VLAN 3内以广播方式发送一个ARP请求帧，查寻主机C的MAC地址。

（8）主机C在收到ARP广播帧后向Port1.2子接口返回一个ARP应答帧。

（9）Port1.2子接口在收到主机C的ARP应答帧后，再次进行数据帧封装，把帧中的目的MAC地址改为主机C的真实MAC地址（其他不变），然后就把主机A发来的数据帧发送给主机 C。这样主机 A 之后要发给 C 的数据帧都先发送给网关，由网关——Port1.2子接口做三层转发。

主机C与主机A之间的通信原理一样，最终实现VLAN间的三层互通。

3．VLAN Switch方案

通过VLAN Switch（VLAN交换）也可以实现不同VLAN间的通信。VLAN交换是一种按照VLAN标签进行数据转发的技术，需要预先在网络中的各交换机上建立一条静态转发路径。当交换机接收到符合转发条件的 VLAN 数据后，根据 VLAN 交换表将报文直接转发到相应的出接口，无需查看 MAC 地址表，提高了转发效率及安全性，可有效地避免MAC地址攻击及广播风暴。

该方案仅在**S7700**、**S9300**、**S9300E**和**S9700**等华为高端**S**系列交换机中支持。

VLAN交换功能如下。

（1）添加外层VLAN标签功能，即VLAN Switch stack-vlan功能。

（2）在不同接口之间转换外层VLAN标签，即VLAN Switch switch-vlan功能。

VLAN Switch stack-vlan功能与VLAN Stacking（VLAN堆叠，将在下章介绍）功能类似，也是一种针对用户不同VLAN封装外层VLAN标签的二层技术。与VLAN Stacking功能的差异如表6-11所示。

表6-11 VLAN Switch功能与VLAN Stacking功能比较

功能	共同点	不同点	优缺点
VLAN Switch stack-vlan	（1）接口在收到的帧的最外层VLAN标签外，再加上一层 VLAN标签 （2）端口处理帧的方式一样，即端口可以配置多个VLAN，端口可以给不同 VLAN的帧加上不同的外层标签；端口在接收帧时，给帧加上外层标签；发送帧时，剥掉帧最外层的标签	VLAN Switch 功能需要预先在网络中各交换节点上建立一条静态转发路径。交换节点接收到符合转发条件的VLAN报文后，根据VLAN Switch表将报文直接转发到相应的接口，无需查看MAC地址表 当VLAN Switch中的任意VLAN与全局VLAN冲突，如果该VLAN已经应用到VLAN Switch功能中，那么全局中将无法创建该VLAN	其优点是报文转发时无需查看MAC地址表，提高了转发效率及安全性，可有效地避免 MAC 地址攻击及广播风暴；其缺点是如果有大量的用户接入交换节点，对每一个用户都需要进行初始配置，建立静态转发路径。使得网络管理者的任务量加大，不利于管理
VLAN Stacking		配置VLAN Stacking功能后，报文的转发需要依赖MAC地址表	其优点是用户接入方便，不需要网络管理者进行初始配置。通过MAC地址表指导报文转发；其缺点是报文的转发效率低，易产生广播风暴和受到 MAC地址攻击

VLAN Switch switch-vlan功能与VLAN Mapping（VLAN映射，将在下章具体介绍）功能类似，也可以实现不同VLAN间的通信。与VLAN Mapping功能的差异如表 6-12所示。

表6-12 VLAN Switch功能与VLAN Mapping功能比较

功能	共同点	不同点	优缺点
VLAN Switch switch-vlan	（1）接口在收到带有 VLAN 标签帧后，对外层 VLAN 标签进行替换操作 （2）当在端口上配置了 VLAN Switch 或 VLAN Mapping 功能后，端口在向外发送本地 VLAN 帧时，将帧中的 VLAN 标签替换成外部 VLAN 的 VLAN 标签 （3）当端口在接收外部 VLAN 帧时，将帧中的 VLAN 标签替换成本地 VLAN 的 VLAN 标签	VLAN Switch 功能需要预先在网络中各交换节点上建立一条静态转发路径。交换节点接收到符合转发条件的 VLAN 报文后，根据 VLAN Switch 表将报文直接转发到相应的接口，无需查看 MAC 地址表 当 VLAN Switch 中的任意 VLAN 与全局 VLAN 冲突，如果该 VLAN 已经应用到 VLAN Switch 功能中，那么全局中将无法创建该 VLAN	其优点是报文转发时无需查看 MAC 地址表，提高了转发效率及安全性，可有效地避免 MAC 地址攻击及广播风暴；其缺点是如果有大量的用户接入交换节点，对每一个用户都需要进行初始配置，建立静态转发路径。使得网络管理者的任务量加大，不利于管理
VLAN Mapping		（1）配置 VLAN Mapping 功能后，报文的转发需要依赖 MAC 地址表 （2）VLAN Mapping 实现两个 VLAN 内设备互相通信时，两个 VLAN 内设备的 IP 地址必须处于同一网段	其优点是用户接入方便，不需要网络管理者进行初始配置。通过 MAC 地址表指导报文转发；其缺点是报文的转发效率低，易产生广播风暴和受到 MAC 地址攻击

以上所介绍的三种 VLAN 间通信方案有各自的优缺点，适用的网络环境也不尽相同，表6-13列出了它们之间的基本特性比较，用户可根据实际组网环境选择合适的方案部署。

表6-13 三种VLAN间通信方案的比较

VLAN 间通信方案	优点	缺点	适用场景
三层 VLANIF 接口方案	属于不同 VLAN 且位于不同网段的用户，只要在路由可达的前提下，随时可以互相通信	如果网络中存在多个用户属于不同 VLAN，那么需要为每一个 VLAN 创建对应的 VLANIF 接口，并分配 IP 地址，增加了配置工作量且占用大量 IP 地址资源	适用于规模小、IP 地址固定的网络，且用户属于不同网段 如果 VLAN 配置比较多，既要进行二层转发，又要进行三层转发，选择使用 VLANIF 接口
三层以太网子接口方案	属于不同 VLAN 且位于不同网段的用户，只要在路由可达的前提下，随时可以互相通信	如果网络中存在多个用户属于不同 VLAN，那么需要在设备上为每一个 VLAN 创建一个子接口，并分配 IP 地址，增加了配置工作量且占用大量 IP 地址资源	适用于规模小的网络，且用户属于不同网段。如果主要以三层转发为主，选择使用子接口
VLAN Switch 方案	报文转发时无需查看 MAC 地址表，提高了转发效率及安全性	如果有大量的用户接入交换机，需要对每一个用户进行初始配置，建立静态转发路径，使得网络管理者的任务量加大	适用于规模小、网络环境较固定的场景中

6.9.3 配置通过VLANIF接口实现VLAN间通信

划分VLAN后，同一VLAN内的用户可以互相通信，但是属于不同VLAN的用户 不能直接通信。为了实现VLAN间通信，可通过配置逻辑的三层接口——VLANIF接口来实现。

VLANIF是逻辑三层口，配置IP地址后可实现网络层互通。通过VLANIF接口实现VLAN间通信需要为每个VLAN创建对应的逻辑接口VLANIF接口，并为每个VLANIF接口配置IP地址实现三层互通。另外，为了成功实现VLAN间的三层互通，VLAN内用户主机的缺省网关必须对应VLANIF接口的IP地址。

配置通过VLANIF接口实现VLAN间通信的配置步骤如表6-14所示。

表6-14 通过VLANIF接口实现VLAN间通信的配置步骤

步骤	命令	说明
1	**system-view** 例如：<HUAWEI> system-view	进入系统视图
2	**interface vlanif** *vlan-id* 例如：[HUAWEI] **interface vlanif** 10	进入 VLANIF 接口视图。参数用来指定要配置 VLANIF 接口的 VLAN ID，取值范围为 1～4 094 整数。但 VLANIF 接口的编号必须对应一个已创建的 VLAN。且只有当 VLAN 内存在（至少一个）状态为 Up 的物理端口时，该 VLAN 对应的 VLANIF 接口状态才会 Up
3	**ip address** *ip-address* { *mask* | *mask-length* }[**sub**] 例如：[HUAWEI –Vlanif10] **ip address 10.1.1.2 8**	为以上 VLANIF 接口配置主或从 IP 地址，以实现 VLAN 三层互通。命令中的参数说明如下。 （1）*ip-address*：指定 VLANIF 接口的 IPv4 地址，为点分十进制格式 （2）*mask*：二选一参数，指定以上配置的 IP 地址所对应的子网掩码，也为点分十进制格式 （3）*mask-length*：二选一参数，指定以上配置的 IP 地址所对应的子网掩码前缀长度，为 1～32 的整数 （4）**sub**：可选项，指定所配置的 IP 地址为从 IP 地址，不选择此可选项，则所配置的 IP 地址为主 IP 地址。有时为了使交换机的一个接口能够与多个子网相连，可以在一个接口上配置多个 IP 地址。主 IP 地址只能配置一个，但可以配置多个从 IP 地址。当配置主 IP 地址时，如果接口上已经有主 IP 地址，则原主 IP 地址被删除，新配置的 IP 地址成为主 IP 地址 可用 **undo ip address** [*ip-address* { *mask* | *mask-length* } [**sub**]]命令删除 VLANIF 接口上配置的指定 IP 地址
4	**damping time** *delay-time* 例如：[HUAWEI-Vlanif10] **damping time 10**	（可选）配置 VLAN Damping 功能的抑制时间。参数 *delay-time* 用来指定 VLANIF 变为 Down 的延迟时间，取值范围为 0～20 的整数秒 【说明】为避免因 VLANIF 接口状态变化引起的网络震荡，可在 VLANIF 接口上通过本命令使能 VLAN Damping 功能。这时，当 VLAN 中最后一个处于 Up 状态的成员端口变为 Down 后，启动 VLAN Damping 功能的设备会抑制设定的时间后再上报给 VLANIF 接口。如果在抑制的时间内 VLAN 中有成员口状态变为 Up，则 VLANIF 接口状态保持 Up 不变 缺省情况下，抑制时间是 0 秒，表示去使能 VLAN Damping 功能，可用 **undo damping time** 命令恢复 VLANIF 变为 Down 的延迟时间为 0 秒

（续表）

步骤	命令	说明
5	**mtu** *mtu* 例如：[HUAWEI-Vlanif10] **mtu 1492**	（可选）配置 VLANIF 接口的 MTU（Maximum Transmission Unit，最大传输单元），取值范围为 128～9216 的整数字节 缺省情况下，MTU 取值为 1500 字节，可用 **undo mtu** 命令恢复 VLANIF 接口的最大传输单元为缺省值
6	**bandwidth** *bandwidth* 例如：[HUAWEI-Vlanif10] **bandwidth** 1000	（可选）配置 VLANIF 接口的带宽，取值范围是 1～1000000 的整数 Mbit/s。配置 VLANIF 接口的带宽用于网管获取带宽，便于监控流量

【示例 1】配置VLANIF2接口的主IP地址为10.1.1.2，从IP地址为11.1.1.3，子网掩码都为255.0.0.0。

```
<HUAWEI>system-view
[HUAWEI] interface vlanif 2
[HUAWEI-Vlanif2] ip address 10.1.1.2 8
[HUAWEI-Vlanif2] ip address 11.1.1.3 255.0.0.0 sub
```

【示例 2】配置 VLAN 10 向 VLANIF10 上报 Down 的延迟时间为 10s，并配置VLANIF10的最大传输单元为1492个字节，带宽为10000Mbit/s。

```
<HUAWEI> system-view
[HUAWEI] vlan 10
[HUAWEI-vlan10] quit
```

[HUAWEI] interface vlanif10

[HUAWEI-Vlanif10] damping time10

[HUAWEI-Vlanif10] mtu 1492

[HUAWEI-Vlanif10] bandwidth 10000

6.9.4 通过VLANIF接口实现VLAN间通信的配置示例

本示例拓扑结构如图6-22所示。企业的不同用户拥有相同的业务，但位于不同的网段，而相同业务的用户又属不同 VLAN，现需要实现不同VLAN中的用户相互通信。如User1和User2中拥有相同的业务，但是属于不同的VLAN且位于不同的网段。现需要实现User1和User2互通。

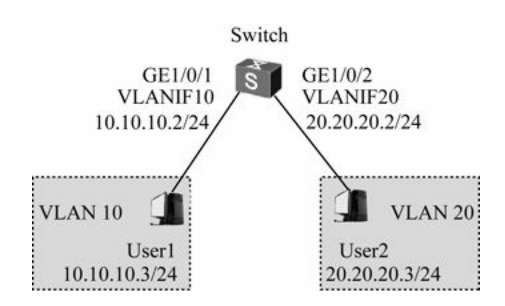


图6-22 通过VLANIF接口实现VLAN间通信的示例

本示例很简单，基本配置思路如下。

（1）创建VLAN，确定用户所属的VLAN。

（2）配置端口加入 VLAN，允许用户所属的VLAN通过当前端口。

（3）创建VLANIF接口并配置IP地址，利用三层交换机的IP路由功能即可实现三层互通。

为了实现VLAN间互通，VLAN内主机的缺省网关必须配置为对应的VLANIF接口的IP地址。具体配置步骤如下。

1）批量创建VLAN 10和VLAN 20。

<HUAWEI>system-view

[HUAWEI] vlan batch 10 20

2）把User1和User2所连接的交换机端口分别加入对应的VLAN中（端口类型均 配置为Access类型）。

[HUAWEI] interface gigabitethernet 1/0/1

[HUAWEI-GigabitEthernet1/0/1] port link-type access

[HUAWEI-GigabitEthernet1/0/1] port default vlan 10

[HUAWEI-GigabitEthernet1/0/1] quit

[HUAWEI] interface gigabitethernet 1/0/2

[HUAWEI-GigabitEthernet1/0/2] port link-type access

[HUAWEI-GigabitEthernet1/0/2] port default vlan 20

[HUAWEI-GigabitEthernet1/0/2] quit

3）为VLAN 10和VLAN 20分别配置VLANIF接口 IP地址。

[HUAWEI] interface vlanif10

[HUAWEI-Vlanif10] ip address 10.10.10.2 24

[HUAWEI-Vlanif10] quit

[HUAWEI] interface vlanif 20

[HUAWEI-Vlanif20] ip address 20.20.20.2 24

[HUAWEI-Vlanif20] quit

4）在VLAN10中的User1主机上配置IP地址为10.10.10.3/24，缺省网关为VLANIF10接口的IP地址10.10.10.2/24；在VLAN20中的User2主机上配置IP地址为20.20.20.3/24，缺省网关为VLANIF20接口的IP地址20.20.20.2/24。

配置完成后，VLAN10内的User1与VLAN20内的User2能够相互访问，通过ping命令即可进行测试。

6.9.5 通过VLANIF接口实现跨越三层网络通信的配置示例

本示例拓扑结构如图6-23所示。示例中互联的三层交换机SwitchA和SwitchB的下面都连接了一个VLAN 10的二层网络，要求SwitchA和SwitchB之间通过OSPF路由协议实现两个VLAN 10二层网络中的用户PC的三层互通。

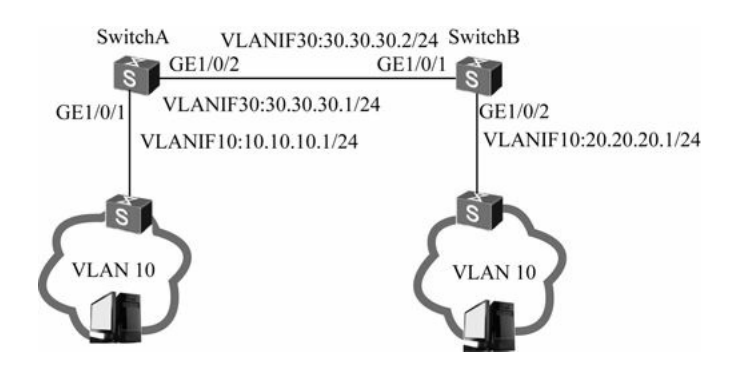


图6-23 通过VLANIF接口跨越三层网络实现VLAN间通信的示例

1．配置思路分析

本示例与上一节介绍的示例在配置上的唯一区别就是要通过OSPF路由协议实现连接在不同网络中的不同VLAN间三层互通。但这里要特别注意的一点是，目前在华为交换机中仍不能直接在物理接口上配置IP地址，只能通过把物理接口放进一个VLAN中，然后为VLAN配置VLANIF接口IP地址来实现物理接口的三层化。

在图中已标识了各VLANIF接口的 IP地址，这里要注意，虽然同为VLAN 10，但因为不是一个网段中的VLAN，所以是不同的，可以为它们的VLANIF接口配置不同网段的IP地址。通过OSPF协议对10.10.10.0/24、20.20.20.0/24和30.30.30.0/24三个子网 的通告，就相当于通过OSPF协议实现这三个子网间的路由通信。

2．配置步骤

SwitchA上的配置：

（1）创建VLAN10和VLAN30。

<HUAWEI>system-view

[HUAWEI] sysname SwitchA

[SwitchA] vlan batch 10 30

（2）把GE1/0/1和GE1/0/2端口均配置为Trunk类型，分别加入VLAN 10和VLAN 30中。

[SwitchA] interface gigabitethernet 1/0/1

[SwitchA-GigabitEthernet1/0/1] port link-type trunk

[SwitchA-GigabitEthernet1/0/1] port trunk allow-pass vlan 10　!---允许VLAN 10通过

[SwitchA-GigabitEthernet1/0/1] quit

[SwitchA] interfacegigabitethernet 1/0/2

[SwitchA-GigabitEthernet1/0/2] port link-type trunk

[SwitchA-GigabitEthernet1/0/2] port trunk allow-pass vlan 30　!---允许VLAN 30通过

[SwitchA-GigabitEthernet1/0/2] quit

【经验之谈】如果你是一位细心的读者，你就可能在看到以上配置时产生一个疑问，为什么在Gigabitethernet 1/0/2端口上不需要同时允许VLAN 10呢?其实最根本的原因还是因为华为S系列交换机的物理端口不能直接配置IP地址，是二层接口，必须通过加入一个VLAN，然后通过对应的VLANIF接口来实现二层与三层之间的转换，配置IP地址，进行三层通信。这里的VLANIF接口其实就相当于路由器上的路由接口。至于接口类型，其实无所谓，因为在三层通信中，数据到达交换机后，三层模块会去掉数据链路层协议头，包括VLAN标签部分。另外要注意的是，在华为S系列交换机中同一交换机中各VLAN缺省都是三层互通的。

（3）配置VLANIF10和VLANIF30的IP地址分别为10.10.10.1/24和30.30.30.1/24。

[SwitchA] interface vlanif 10

[SwitchA-Vlanif10] ip address 10.10.10.1 24

[SwitchA-Vlanif10] quit

[SwitchA] interface vlanif 30

[SwitchA-Vlanif30] ip address 30.30.30.1 24

[SwitchA-Vlanif30] quit

（4）配置OSPF路由，宣告VLAN 10和VLAN 30所对应的网段。有关OSPF路由的配置请参见《华为路由器学习指南》。

[SwitchA] router id 1.1.1.1　!----配置路由器 ID（可随意，通常是三层设备的 lookback0接口 IP地址）

[SwitchA] ospf

[SwitchA-ospf-1] area 0

[SwitchA-ospf-1-area-0.0.0.0] network10.10.10.0 0.0.0.255　!----宣告VLAN 10所在的网段

[SwitchA-ospf-1-area-0.0.0.0] network30.30.30.0 0.0.0.255　!----宣告VLAN 30所在的网段

[SwitchA-ospf-1-area-0.0.0.0] quit

SwitchB上的配置：

（1）创建VLAN10和VLAN30。

<HUAWEI>system-view

[HUAWEI] sysname SwitchB

[SwitchB] vlan batch 10 30

（2）把GE1/0/2和GE1/0/1端口均配置为Trunk类型，分别加入VLAN 10和VLAN 30中。

[SwitchB] interface gigabitethernet 1/0/2

[SwitchB-GigabitEthernet1/0/2] port link-type trunk

[SwitchB-GigabitEthernet1/0/2] port trunk allow-pass vlan 10

[SwitchB-GigabitEthernet1/0/2] quit

[SwitchB] interface gigabitethernet 1/0/1

```
[SwitchB-GigabitEthernet1/0/1] port link-type trunk
[SwitchB-GigabitEthernet1/0/1] port trunk allow-pass vlan 30
[SwitchB-GigabitEthernet1/0/1] quit
```

（3）配置VLANIF10和VLANIF30的IP地址分别为20.20.20.1/24和30.30.30.2/24。

```
[SwitchB] interface vlanif 10
[SwitchB-Vlanif10] ip address 20.20.20.1 24
[SwitchB-Vlanif10] quit
[SwitchB] interface vlanif30
[SwitchB-Vlanif30] ip address 30.30.30.2 24
[SwitchB-Vlanif30] quit
```

（4）配置OSPF路由，宣告VLAN 10和VLAN 30所对应的网段。

```
[SwitchB] router id 2.2.2.2
[SwitchB] ospf
[SwitchB-ospf-1] area 0
[SwitchB-ospf-1-area-0.0.0.0] network20.20.20.0 0.0.0.255
[SwitchB-ospf-1-area-0.0.0.0] network30.30.30.0 0.0.0.255
[SwitchB-ospf-1-area-0.0.0.0] quit
```

另外，需要把SwitchA下挂的二层网络中PC上配置缺省网关为VLANIF10接口的IP地址10.10.10.1/24；把SwitchB下挂的二层网络中PC上配置缺省网关为VLANIF10接口的IP地址20.20.20.1/24。

配置完成后，两个二层网络的PC就可实现二层隔离，但三层互通。

6.9.6 配置通过子接口实现VLAN间通信

属于不同 VLAN 且位于不同网段的用户，可通过部署子接口、配置 IP 地址并与VLAN相关联，通过三层网络实现VLAN间通信。同样，为了成功实现VLAN间互通， VLAN内主机的缺省网关必须是对应子接口的IP地址。具体的配置步骤如表6-15所示（仅**S5700HI**和**5710EI**子系列、**S7700**、**S9300**和**S9700**系列华为交换机支持 ）。

表6-15 通过子接口实现VLAN间通信的配置步骤

步骤	命令	说明
1	**system-view** 例如：< HUAWEI > **system-view**	进入系统视图
2	**Interface** { **ethernet** | **gigabitethernet** | **xgigabitethernet** | **eth-trunk** } *interface-number.subinterface-number* 例如：[HUAWEI]**interface** gigabitethernet 0/0/1.1	键入要配置的三层以太网子接口，进入子接口视图
3	**ip address** *ip-address* { *mask* | *mask-length* } [**sub**] [HUAWEI-GigabitEthernet0/0/1.1] **ip address** 192.168.10.1 255.255.255.0	为以上子接口配置主、从 IP 地址（**但交换机上各接口上配置的所有 IP 地址不能位于相同子网**），实现三层互通。命令中的参数说明如下。 （1）*ip-address*：指定 VLANIF 接口的 IPv4 地址，为点分十进制格式 （2）*mask*：二选一参数，指定以上配置的 IP 地址所对应的子网掩码，也为点分十进制格式 （3）*mask-length*：二选一参数，指定以上配置的 IP 地址所对应的子网掩码前缀长度，为 1～32 的整数

（续表）

步骤	命令	说明
3	**ip address** *ip-address* { *mask* | *mask-length* } [**sub**] [HUAWEI-GigabitEthernet0/0/1.1] **ip address** 192.168.10.1 255.255.255.0	（4）**sub**：可选项，指定所配置的 IP 地址为从 IP 地址，不选择此可选项，则所配置的 IP 地址为主 IP 地址。有时为了使交换机的一个接口能够与多个子网相连，可以在一个接口上配置多个不同子网中的 IP 地址。当配置主 IP 地址时，如果接口上已经有主 IP 地址，则原主 IP 地址被删除，新配置的 IP 地址成为主 IP 地址 缺省情况下，没有配置 IP 地址，可用 **undo ip address** [*ip-address* { *mask* | *mask-length* } [**sub**]]命令删除子接口上配置的指定 IP 地址
4	**dot1q termination vid** *low-pe-vid* [**to** *high-pe-vid*] 例如：[HUAWEI -GigabitEthernet0/0/1.1] **dot1q termination vid 100**	配置子接口 dot1q 封装的单层 VLAN ID。命令中的参数说明如下。 （1）*low-pe-vid*：指定用户数据帧中的 VLAN 标签的取值下限，取值范围是 2～4 094 的整数 （2）*high-pe-vid*：可选参数，指定用户数据帧中的 VLAN 标签的取值上限，取值范围是 2～4 094 的整数 【说明】当子接口用于三层转发时，不支持将通过的 VLAN 配置成一段。不同主接口下的子接口可以关联相同的 VLAN ID，但是同一主接口下的不同子接口一定不能关联相同的 VLAN ID 缺省情况，子接口没有配置 dot1q 封装的单层 VLAN ID，可用 **undo dot1q termination vid** *low-pe-vid* [**to** *high-pe-vid*]命令取消子接口 dot1q 封装的单层 VLAN ID
5	**arp broadcast enable** 例如：[HUAWEI -GigabitEthernet0/0/1.1] **arp broadcast enable**	使能子接口的 ARP 广播功能。当 IP 数据帧需要从终结子接口发出，但是没有相应的 ARP 表项时： （1）如果接入设备能够主动发送 ARP 数据帧，则不需要配置终结子接口的 ARP 广播功能，就可以实现从该终结子接口的转发。 （2）如果接入设备不能够主动发送 ARP 数据帧，但终结子接口上未使能 ARP 广播功能，那么系统会直接把该 IP 数据帧丢弃。此时该终结子接口的路由可以看作是黑洞路由。 （3）如果接入设备不能够主动发送 ARP 数据帧，但终结子接口上已使能 ARP 广播功能，那么系统会构造带 Tag 的 ARP 广播数据帧，然后从该终结子接口发出。 缺省情况下，终结子接口没有使能 ARP 广播功能，可用 **undo arp broadcast enable** 命令去使能终结子接口的 ARP 广播功能

【示例 】在GE0/0/1.1子接口上配置封装方式为dot1q，通过的报文外层VLAN标签为10，并使能该子接口的ARP广播功能。

```
<HUAWEI>system-view
[HUAWEI] interface gigabitethernet 0/0/1.1
[HUAWEI-GigabitEthernet0/0/1.1] dot1q termination vid 10
[HUAWEI-GigabitEthernet0/0/1.1] arp broadcast enable
```

6.9.7 通过子接口实现VLAN间通信的配置示例

本示例拓扑结构如图 6-24 所示，企业的不同部门拥有相同的业务，如上网、VoIP 等业务，且各个部门中的用户位于不同的网段。因存在不同的部门中相同的业务所属的VLAN各不相同的现象，现需要实现相同业务的不同 VLAN 中的用户相互通信。如部门 1 （ Department1 ）和部门 2 （Department2）中拥有相同的上网业务，但是属于不同的VLAN且位于不同的网段。现需要实现部门1与部门2的用户互通。

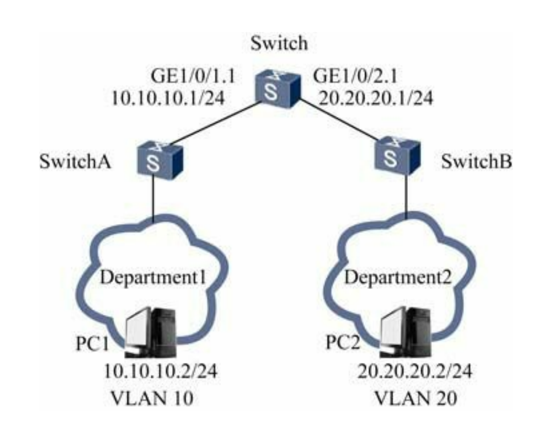


图6-24 通过子接口实现VLAN间通信的示例

1．配置思路分析

本示例很显然子接口必须在Switch上配置，因为只有它同时连接了VLAN 10和 VLAN 20。在此仅介绍与通过子接口VLAN间通信的相关配置（可直接按照表 6-23所示的步骤分别为VLAN 10和VLAN 20创建对应的子接口），至于 SwitchA和 SwitchB各Trunk接口的配置此处不作介绍。

2．配置步骤

此处仅介绍Switch上的相关配置。

（1）在 Switch的GE1/0/1端口上创建用于连接 SwitchA下面的VLAN 10用户的三层以太网子接口GE1/0/1.1，并封装VLAN 10，配置与VLAN 10所在网段相同的IP地址。

```
<HUAWEI>system-view
[HUAWEI] interfacegigabitethernet1/0/1.1
[HUAWEI-GigabitEthernet1/0/1.1] dot1q termination vid 10
[HUAWEI-GigabitEthernet1/0/1.1] ip address 10.10.10.1 24
[HUAWEI-GigabitEthernet1/0/1.1] arp broadcast enable
[HUAWEI-GigabitEthernet1/0/1.1] quit
```

（2）在 Switch的GE1/0/2端口上创建用于连接 SwitchB下面的VLAN 20用户的三层以太网子接口GE1/0/2.1，并封装VLAN 20，配置与VLAN 20所在网段相同的IP地址。

```
[HUAWEI] interface gigabitethernet 1/0/2.1
[HUAWEI-GigabitEthernet1/0/2.1] dot1q termination vid 20
[HUAWEI-GigabitEthernet1/0/2.1] ip address 20.20.20.1 24
[HUAWEI-GigabitEthernet1/0/2.1] arp broadcast enable
[HUAWEI-GigabitEthernet1/0/2.1] quit
```

另外，需要在 VLAN 10 中的 PC1 上配置缺省网关为 GE1/0/1.1接口的 IP地址10.10.10.1/24；在 VLAN 20 中的 PC2 上配置缺省网关为 GE1/0/2.1 接口的 IP 地址20.20.20.1/24。

配置完成后，VLAN 10内的PC1与VLAN 20内的PC2就能够实现三层互通了。

6.9.8 配置通过VLAN Switch实现VLAN间通信

VLAN Switch是一种按照VLAN标签进行数据转发的技术，需要预先在网络中的各交换机上建立一条静态转发路径。当交换机某端口在接收到符合转发条件的VLAN数据后，根据VLAN Switch表中建立的外层VLAN映射关系将数据帧替换外层VLAN标签 后从指定端口发送出去，无需查看 MAC 地址表，提高了转发

效率及安全性，同时还可有效地避免 MAC 地址攻击及广播风暴。在规模很小、较固定的网络环境中，可以配置VLAN Switch实现不同VLAN间互通。

VLAN Switch功能仅在 **SS7700**、**S9300**、**S9300E**和 **S9700**等华为高端 **S**系列交换机中支持，可通过以下两种方式实现VLAN间通信。

（1）替换外层VLAN标签，即VLAN Switch switch-vlan功能，与VLAN Mapping （VLAN映射）功能类似，可实现VLAN间通信。本节主要介绍VLAN Switch switch-vlan功能的配置方法。

（2）添加外层VLAN标签功能，即VLAN Switch stack-vlan功能，与VLAN Stacking功能类似，是一种针对用户不同VLAN封装外层VLAN标签的二层技术。关于VLAN Switch stack-vlan功能的配置将在本章后面介绍QinQ协议时再介绍。

通过VLAN Switch switch-vlan功能实现VLAN间通信的配置步骤很简单，只需在系统视图下配置 vlan-switch vlan-switch-name interface interface-type1 interface-number1 vlan vlan-id1 [inner-vlan vlan-id2 [to vlan-id3]] interface interface-type2 interface-number2 [switch-vlan vlan-id4] 命令，配置VLAN Switch switch-vlan功能，替换外层VLAN标签。命令中的参数说明如下。

（1）vlan-switch-name：指定要配置的VLAN Switch的名称，长度为 1～32的字符串。

（2）interface-type1 interface-number：指定要替换外层 VLAN 标签的交换机端口，即源端口。

（3）interface-type2 interface-number：指定替换了外层VLAN标签后VLAN帧的发出端口，即目的端口。

（4）vlan-id1：指定替换前的外层VLAN ID。

（5）vlan-id2 to vlan-id3：可选参数，指定替换前的内层VLAN ID或一个VLAN ID范围。

（6）vlan-id4：可选参数，指定替换后的外层VLAN ID。

以上 vlan id参数的取值范围均为 2～4 094，但要注意，VLAN Switch中要替换成的VLAN不能是全局VLAN，也就是不能在整个网络中都存在VLAN，只能在单个设备中存在VLAN。可用undo vlan-switch vlan-switch-name命令删除对应的VLAN Switch配置。

【示例 】配置VLAN Switch的转换外层VLAN 标签功能，将GE1/0/1端口接收到的带有VLAN 10标签的报文变换成带有VLAN 20标签的报文从GE1/0/2端口转发出去。

<HUAWEI>system-view

[HUAWEI] vlan-switch name1 interfacegigabitethernet1/0/1 vlan 10 interfacegigabitethernet 1/0/2 switch-vlan 20

6.9.9 通过VLAN Switch实现VLAN间通信的配置示例

本示例拓扑结构如图6-25所示。Switch的接口GE1/0/1、GE1/0/2分别与SwitchA、SwitchB上行接口相连，SwitchA、SwitchB的下行接口分别加入VLAN 10、VLAN 20。现要通过VLAN Switch功能实现VLAN 10内的PC与VLAN 20内的PC能够互访。

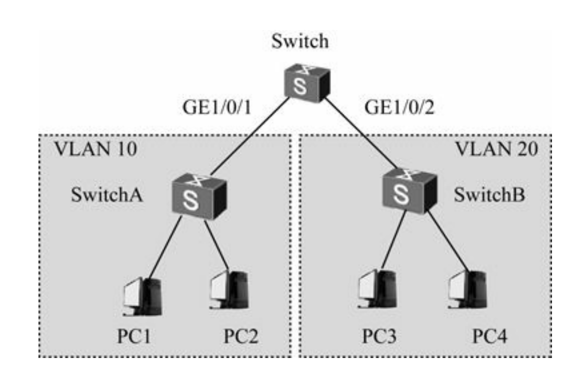


图6-25 通过VLAN SwitchVLAN间通信的配置示例

配置方法如下。

（1）配置SwitchA、SwitchB的相关接口加入VLAN（略）。

（2）配置Switch的VLAN Switch功能。把GE1/0/1端口上接收到的VLAN 10帧中的VLAN标签替换成VLAN 20，然后从GE1/0/2端口发送出去。

<HUAWEI>system-view

[HUAWEI] vlan-switch name1 interfacegigabitethernet1/0/1 vlan 10 interfacegigabitethernet1/0/2 switch-vlan 20

注意

VLAN10、VLAN20不能是Switch已经创建的全局VLAN，且GE1/0/1端口不能 加入VLAN 10，GE1/0/2端口不能加入VLAN 20，否则无法配置VLAN Switch。

配置完成后，VLAN10内的PC与VLAN20内的PC能够相互访问。从中看出，采用VLAN Switch功能实现VLAN间通信的配置很简单。

6.9.10 VLAN间通信配置管理

配置好以上介绍的各种VLAN间通信方案后，可在任意视图下执行以下display命令查看相关的配置信息，以验证配置效果。

（1）使用**display vlan** [vlan-id [**verbose**]] 命令查看所有VLAN或指定VLAN的显示信息。

（2）使用**display interface vlanif** [vlan-id] 命令查看所有或指定VLANIF接口的状态信息、配置信息和统计信息。用户可以根据这些信息进行接口的故障诊断等。

（3）使用 **display vlan-switch** [vlan-switch-name | **interface** interface-type interface- number] 命令查看所有或指定的 VLAN Switch 的配置信息。可查看的信息包括 VLAN Switch名称、源接口编号、源VLAN编号、目的接口编号、目的VLAN编号、操作类型、VLAN Switch当前状态。

6.10 管理VLAN的配置与管理

在华为的S系列交换机中，物理接口都是不能直接配置IP地址的，在进行交换机的远程管理（如通过远程Telnet方式访问）中，都是利用VLANIF逻辑接口（或者其他逻辑接口）上配置的IP地址作为管理IP地址。而这个管理IP地址所对应的VLAN称为“管理VLAN”。

一般情况下，任意 VLAN 都可成为管理 VLAN（当然，这是指三层交换机上的VLAN），但为了避免一些非授权进行远程交换机管理的 VLAN 用户对交换机进行非法管理，提高交换机的安全性，在华为 S 系

列交换机中一旦某个 VLAN 被配置为管理VLAN，则不允许Access类型和Dot1q-tunnel类型端口加入该VLAN。

管理VLAN功能部署成功后，用户可通过管理VLAN对应的VLANIF接口的IP地址Telnet到管理交换机，从而实现通过远端设备集中管理。管理VLAN功能的具体配置步骤如表6-16所示。

表6-16 管理VLAN的配置步骤

步骤	命令	说明
1	**system-view** 例如：< HUAWEI > **system-view**	进入系统视图
2	**vlan** *vlan-id* 例如：[HUAWEI] **vlan** 5	创建并进入要配置为管理 VLAN 的 VLAN 视图
3	**management-vlan** 例如：[HUAWEI-vlan5] **management-vlan**	把以上 VLAN 配置为管理 VLAN。使用 **management-vlan** 命令配置 VLAN 为管理 VLAN 后，不允许 Access 类型和 Dot1q-tunnel 类型接口加入该 VLAN，**而只能是 trunk 或 hybrid 类型的端口** 缺省情况下，该 VLAN 没有配置为管理 VLAN，可用 **undo management-vlan** 命令取消配置 VLAN 为管理 VLAN
4	**quit** 例如：[HUAWEI-vlan5] **quit**	退出 VLAN 视图，返回系统视图
5	**interface vlanif** *vlan-id* 例如：[HUAWEI] **interface vlanif** 5	创建以上管理 VLAN 的 VLANIF 接口，并进入 VLANIF 接口视图
6	**ip address** *ip-address* { *mask* | *mask-length* } [**sub**] 例如：[HUAWEI-Vlanif5] **ip address 10.1.1.2 8**	为以上管理 VLAN 的 VLANIF 接口配置主或从 IP 地址。有关参数和其他说明参见表 6-15 的第 3 步

管理VLAN对应的VLANIF接口IP地址配置成功后，可通过命令telnet登录到管理交换机实现网管集中管理设备。管理VLAN配置成功后，可以通过display vlan命令查看管理VLAN的配置信息，带有*的VLAN为管理VLAN。

第7章　扩展VLAN特性配置与管理

第6章我们介绍了VLAN划分、VLAN注册、VLAN间通信、管理VLAN等基本特性的配置与管理。本章接着介绍多个非常重要并且在实际的网络设备管理中经常使用的扩展VLAN特性的配置与管理，如VLAN聚合（Super-VLAN）、VLAN内用户隔离（MUX VLAN）、VLAN映射（VLAN Mapping）、VLAN多标签封装（QinQ）等。

以上这些扩展VLAN特性是为了解决在普通VLAN应用过程中出现的一些问题而开发的。如VLAN聚合是为了解决普通VLAN间通信必须使每个VLAN位于不同的IP子网，必须为每个VLAN配置VLANIF接口，并分配IP地址的问题；MUX VLAN是为了达到VLAN内部用户间的二层隔离的目的，同时可实现VLAN间的通信；VLAN映射可解决由ISP公网VLAN到用户私网VLAN的映射问题。

7.1 VLAN聚合配置与管理

由于VLAN本身基本特性的限制，致使在VLAN的使用过程中又遇到了一些问题。如在普通VLAN间通信过程中需要为每个VLAN配置一个VLANIF接口IP地址，同时需要为每个VLAN单独使用一个IP子网，这样就会导致整个公司网络IP子网数可能非常多，最终也将导致 IP 地址浪费的现象也非常严重。为了解决这一问题，就诞生了一种可以聚合多个不配置VLANIF接口的超级VLAN（Super-VLAN）技术，即本节将要介绍的VLAN聚合（VLAN Aggregation）技术。这个超级VLAN可以包含多个位于同一IP子网的VLAN，并且只需要使用一个VLANIF接口IP地址作为各成员VLAN的共同网关即可实现同一超级VLAN内不同成员VLAN间，以及与外部网络间的通信。

本项扩展VLAN特性除了S1700、S2700SI系列外的其他所有华为S系列交换机均支持。

7.1.1 普通VLAN部署的不足

在普通的VLAN部署中，一般是采用一个VLAN对应一个三层VLANIF逻辑接口的方式实现VLAN间的互通。这样部署的结果就导致了IP地址的浪费，因为这样部署后每个VLAN都需要使用一个独立的IP子网，而且要为每个VLAN配置一个带有IP地址的VLANIF接口。

如图 7-1所示，在一个三层交换机（L3 Switch）上部署了 3个VLAN（VLAN 2、VLAN 3和VLAN 4），并为它们创建了三层VLANIF接口，各配置了一个 IP地址，以便实现这 3个VLAN间的三层通信。现如果VLAN 2中预计未来有 10个主机地址的需求，则至少要为其分配一个子网掩码长度是28的子网

1.1.1.0/28，同时需要为其配置一个缺省网关地址，即VLANIF2的IP地址（假设为1.1.1.1），这样一来该子网中可以分配给主机使用的 IP地址共 13个，尽管VLAN 2只需要 10个地址。

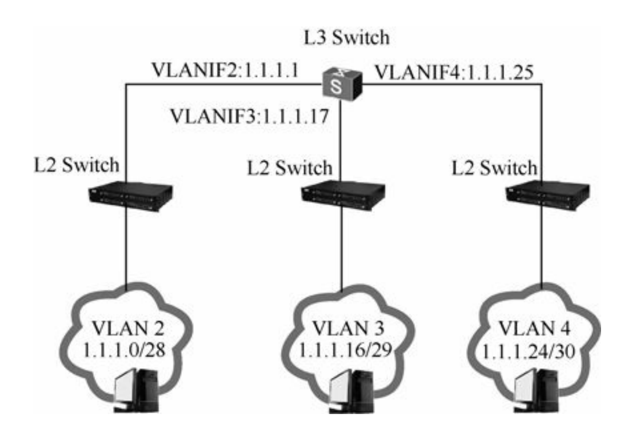


图7-1 普通VLAN配置方式的IP地址分配示例

同理，如果VLAN 3中预计未来有 5个主机地址的需求，至少需要分配一个子网掩码长度是29的子网1.1.1.16/29，也要配置一个缺省网关地址，即VLANIF3的IP地址（假设为 1.1.1.17）。如果VLAN 4中预计未来只有 1个主机，则至少要分配一个子网掩码长度是30的子网1.1.1.24/30，也要配置一个缺省网关地址，即VLANIF4的IP地址（假设 为1.1.1.25）。此时，这三个VLAN所属子网的IP地址分配如表7-1所示。

表7-1 普通VLAN配置方式下的主机IP地址分配示例

VLAN	子网	网关地址	可用地址数	可用主机数	实际需求
2	1.1.1.0/28	1.1.1.1	14	13	10
3	1.1.1.16/29	1.1.1.17	6	5	5
4	1.1.1.24/30	1.1.1.25	2	1	1

从以上介绍可以看出，这三个VLAN一共只需要16（=10+5+1）个主机IP地址，但是按照以上普通VLAN的编址方式，即使最优化的方案也需要占用28（=16＋8＋4）个 IP地址，浪费了将近一半的地址。而且如果VLAN 2后来并没有 10台主机，而实际只接入了3台主机，那么多出来的地址也会因不能再被其他VLAN使用而浪费掉。

另外，这种划分也给后续的网络升级和扩展带来了很大不便。假设VLAN 4今后需要再增加两台主机，但1.1.1.24/30后面的地址已经分配给了其他VLAN，如果又不想改变已经分配的 IP地址，则只能再给VLAN 4的新用户重新分配一个的 29位掩码的子网和一个新的VLAN。这样VLAN4中的客户虽然只有3台主机，但是却被分配在两个子网中，并且也不在同一个VLAN内，不利于网络管理。

综上所述，普通VLAN配置方式下，很多IP地址被子网网络地址、子网定向广播地址、子网缺省网关地址（就是各VLANIF接口IP地址）消耗掉，而不能用于VLAN内的主机。同时，这种地址分配的约束也降低了编址的灵活性，使许多闲置地址也被浪费掉。为了解决这一问题就诞生了本节所要介绍的技术——VLAN Aggregation（VLAN聚合）。

7.1.2 VLAN聚合及优势体现

VLAN聚合技术就是把多个不配置三层VLANIF接口，同处一个IP子网的VLAN （称之为Sub-VLAN）当作一个大的、配置三层VLANIF接口的VLAN（称为Super-VLAN）的成员。这些同一 IP 子网下的多个Sub-VLAN 间可以实现用户的二层隔离，同时这些成员VLAN间又可通过上层的Super-VLAN配置的三层

VLANIF接口IP地址作为缺省网关在各成员VLAN间，以及与网络中其他VLAN间进行通信。

1．VLAN聚合中的两类VLAN

在VLAN聚合中涉及到以下两类VLAN。

（1）Super-VLAN：可以把它看成是一个大的 VLAN，或者说它是 Sub-VLAN 的上层 VLAN。但它与通常意义上的 VLAN 不同，因为它的成员就是下面要介绍的Sub-VLAN，而不是交换机端口（里面不能添加交换机端口），但需要创建三层VLANIF接口（所以**Super-VLAN**只能在三层交换机上创建），并配置IP地址。每个**VLAN**聚合中只能有一个**Super-VLAN**。

（2）Sub-VLAN：它是Super-VLAN的成员，每个**VLAN**聚合中可以有一个或多个**Sub-VLAN**。各Sub-VLAN成员都同处于一个IP子网中，用来对同一IP子网中的不同用户进行二层隔离，但不能创建三层VLANIF接口（可以在二层或三层交换机上创建）。这些Sub-VLAN中的成员就是各用户所连接的交换机端口，但各个Sub-VLAN中的用户网关IP地址都是Super-VLAN的VLANIF接口IP地址，以实现Sub-VLAN成员间，以及与外部网络的三层通信。

一个Super-VLAN可以包含一个或多个Sub-VLAN，它们的关系可以用图7-2来表示（图中的Super-VLAN包括了4个Sub-VLAN）。在同一个Super-VLAN中，无论主机属于哪一个Sub-VLAN，它的IP地址都在Super-VLAN的VLANIF接口IP地址所对应的IP子网内。这样各Sub-VLAN共用同一个三层VLANIF接口，既减少了一部分子网网络地址、子网缺省网关地址和子网定向广播地址的消耗，又实现了不同广播域（也就是各Sub-VLAN）使用同一IP子网地址的目的。消除了子网差异，增加了编址的灵活性，减少了闲置地址浪费。

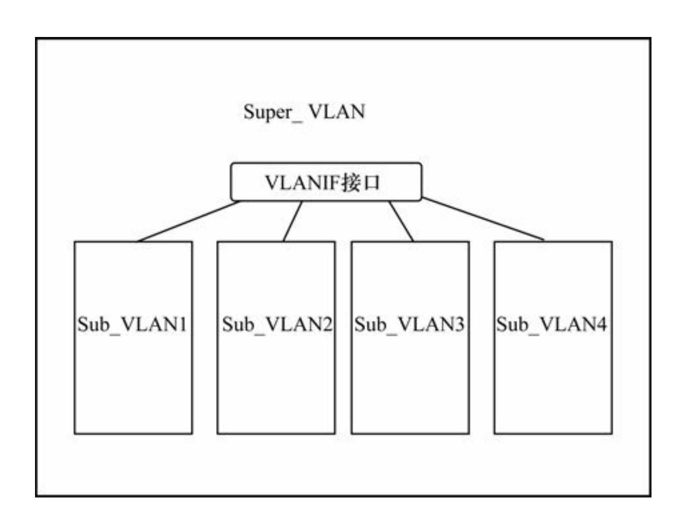


图7-2 Super-VLAN和Sub-VLAN的关系示意图

2．VLAN聚合优势示例

仍以表 7-1所示例子进行说明。假设用户需求不变，仍旧是VLAN 2预计未来有 10个主机地址的需求，VLAN 3预计未来有 5个主机地址的需求，VLAN 4预计未来有 1个主机地址的需求。

按照VLAN聚合的实现方式，新建VLAN 10并配置为Super-VLAN，给其分配一个子网掩码长度是24的子网1.1.1.0/24，并配置其VLANIF接口IP地址为1.1.1.1，如图7-3所示。此时各Sub-VLAN（VLAN2、VLAN3、VLAN4）的IP地址分配如表7-2所示。

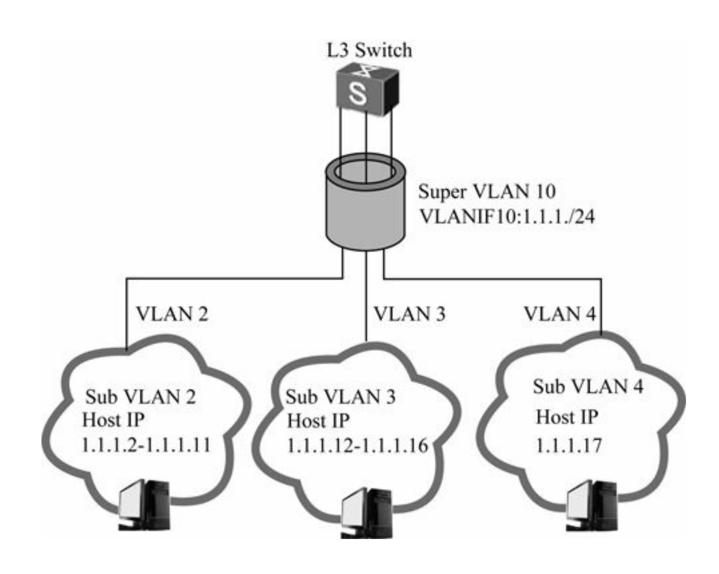


图7-3 VLAN聚合配置方式的 IP地址分配示例

表7-2 VLAN聚合配置方式下的主机 IP地址分配示例

VLAN	子网	网关地址	可用地址数	可用主机数	实际需求
2	1.1.1.0/24	1.1.1.1	10	1.1.1.2～1.1.1.11	10
3			5	1.1.1.12～1.1.1.16	5
4			1	1.1.1.17	1

从表7-2可以看出，在VLAN聚合的实现中，各Sub-VLAN间的界线也不再是从前的子网界线了（因为它们同处一个IP子网中），它们可以根据其各自主机的需求数目在Super-VLAN对应子网内灵活地划分地址范围。另外，VLAN 2、VLAN 3和VLAN 4共用同一个 IP 子网（1.1.1.0/24）、同一个子网缺省网关地址（1.1.1.1）和同一个子网定向广播地址（1.1.1.255）。这样，普通VLAN实现方式中用到的其他子网网络地址（1.1.1.16、1.1.1.24）和子网缺省网关地址（1.1.1.17、1.1.1.25），以及子网定向广播地址（1.1.1.15、1.1.1.23、1.1.1.27）都可以用来作为主机IP地址使用了。也正因如此，以上这三个VLAN一共需要16（=10＋5+1）个IP地址，再加上子网网络地址（1.1.1.0）、子网缺省网关地址（1.1.1.1）和子网定向广播地址（1.1.1.255），一共用去了19个IP地址，该网段内仍剩余255－19＝236的地址可以被任意Sub-VLAN内的主机使用，显得更加灵活，更加实用。

7.1.3 Sub-VLAN通信原理

在VLAN聚合中，Super-VLAN和Sub-VLAN都存在一些特殊性，如Super-VLAN必须配置三层 VLANIF接口，但不能有交换机端口成员；各个 Sub-VLAN 成员同处Super-VLAN的VLANIF接口IP地址所在的一个IP子网中，必须有交换机端口成员，但都不能配置三层的VLANIF接口。正因有以上这些特殊性，造成了Sub-VLAN之间，或者与外部网络间的二、三层通信也存在一定的特殊性。本节要分别予以介绍。

1．Sub-VLAN间的三层通信原理

VLAN 聚合在实现了不同 Sub-VLAN 间共用一个 IP 子网地址的同时也带来了Sub-VLAN间的三层转发问题。因为在普通VLAN实现方式中，VLAN间的主机可以通过各自不同的网关（即各自的VLANIF接口IP地址）进行三层转发来达到互通的目的。但是在VLAN聚合方式下，由于同一个Super-VLAN内的所有主机使用的是同一个IP子网中的IP地址和同一个网关IP地址，即使是属于不同的Sub-VLAN的主机，所以这些主机彼此通信时只会做二层转发，而不会通过网关进行三层转发。而实际上不同的 Sub-VLAN 的主机在二层

是相互隔离的，这就造成了Sub-VLAN间无法二层和三层通信的问题。

解决以上问题的方法就是在作为这些Sub-VLAN网关的Super-VLAN的VLAN接口上启用ARP Proxy（ARP代理）功能，使这个VLANIF接口作为各Sub-VLAN中的主机间通信的ARP代理，代理接收和转发这些主机间的ARP查寻请求和响应包，以最终实现各Sub-VLAN间的三层通信。有关ARP代理的具体知识和配置方法请参见本书第16章相关内容。

如图 7-4所示，在 L3 Switch上创建的 Super-VLAN（VLAN 10）包含 Sub-VLAN （VLAN 2和VLAN 3）。VLAN 2内的主机A与VLAN 3内的主机B的通信过程如下（假设此时主机 A的 ARP映射表中无主机 B的对应表项，且在网关 L3 Switch上使能了Sub-VLAN间的ARP代理功能）。

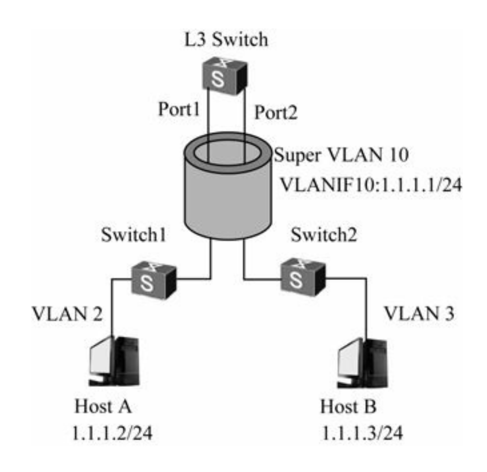


图7-4 通过ARP代理功能实现不同Sub-VLAN间的三层通信示例

（1）主机 A 将主机 B 的 IP 地址（1.1.1.3）和自己所在网段 1.1.1.0/24 进行比较，发现主机 B 和自己在同一个 IP 子网中，但是主机A的ARP映射表中无主机B的对应表项。于是主机A在本VLAN内发送一ARP广播请求包，请求查询主机B的MAC地址。

（2）由于主机B并不在主机A所在的VLAN 2 内，无法接收到主机 A 的这个ARP 请求。但由于在网关（Super-VLAN的VLANIF接口）上使能了Sub-VLAN间的ARP代理功能，所以当网关收到主机A 的ARP请求后，开始在路由表中查找，发现ARP请求中的目的主机B的IP地址（1.1.1.3）为直连接口路由，则网关向所有其他Sub-VLAN中发送一个ARP广播包，代替主机A请求查询主机B的MAC地址。

（3）在各Sub-VLAN中只有IP地址与ARP请求包中的目的IP地址一致的主机才会做出响应。所以在主机B收到网关发送的ARP广播包后，对此请求进行ARP应答。

（4）在网关收到主机B的应答后，建立一个主机B的ARP表项，然后以自己的MAC地址作为源MAC地址（其实是想作为主机B的MAC地址）向主机A发送一个ARP应答包。

（5）在主机A收到来自网关的应答帧后，就以网关MAC地址作为主机B的MAC地址建立主机 B 的 ARP 表项。然后主机 A 便修改要发送给主机 B 的数据帧中的目的MAC地址（目的IP地址不变，仍为主机B的IP地址）为网关的MAC地址（其实，主机A并不知道这个MAC地址不是主机B的MAC地址），把数据发给网关（因为对于主机A来说，它认为主机B的MAC地址就是网关的MAC地址）。

（6）网关在收到主机A发来的数据包后，根据其前面为主机B所创建的ARP映射表项，修改帧中的目的MAC地址为真实的主机B的MAC地址，然后转发数据包到主机B上，实现位于不同Sub-VLAN中的主机间的三层通信。

主机B发送数据帧给主机A的过程与上述的主机A到主机B的数据帧流程类似，不再赘述。

【经验之谈 】 从以上通过ARP代理实现不同VLAN间三层通信的原理可以得出，在每个VLAN中的主机最终只能创建本广播域中的节点（包括网关）ARP映射表项，因为通过网关得到的ARP应答包中的源MAC地址都是网关自己的MAC地址，而不是真正的外部VLAN或网络中主机的MAC地址。只不过通过ARP代理功能可以使网关在与其相连的其他VLAN或网络中代理发送ARP请求，以解析外部VLAN或网络中的主机MAC地址。

2．Sub-VLAN与外部网络的二层通信

我们知道，Super-VLAN与各个Sub-VLAN是作为一个整体与外部网络进行通信的，那么作为Super-VLAN成员的Sub-VLAN在实际的VLAN帧传输中又该如何识别和处理帧中的VLAN标签呢？原来，由于Super-VLAN中没有物理端口成员（也就是没有任何 一个物理端口加入了Super-VLAN），所以在基于端口划分的VLAN（不能像基于MAC地址、IP子网、协议类型和策略的动态VLAN划分的VLAN，因为这些的VLAN中端口成员可动态加入）的二层通信中，无论是数据帧进入交换机端口还是从交换机端口发出都不会有针对Super-VLAN的数据帧。

如图7-5所示，在Switch1上创建了Supuer-VLAN 10，以及VLAN 2和VLAN 3这两个Sub-VLAN。这样从HostA侧Port1进入设备Switch1的帧会被打上VLAN 2的标签，在Switch1中这个标签不会因为VLAN2是VLAN 10的Sub-VLAN而变为VLAN10的标签。该数据帧从Trunk类型的接口Port3出去时，依然是携带VLAN 2的标签。也就是说， Switch1本身不会发出VLAN 10的数据帧。就算其他设备有VLAN 10的数据帧发送到该设备上，这些数据帧也会因为Switch1上没有VLAN 10对应的物理端口成员而最终被丢弃。

但一定要注意，Super-VLAN 中绝对不能存在物理端口的。此时在配置 Trunk 端口和创建Super-VLAN的顺序上还需要注意以下两个方面。

（1）如果先配置了Super-VLAN，再配置Trunk接口时，Trunk的VLAN许可列表项里就自动滤除了Super VLAN。如图 7-5所示，虽然Switch1的Port3允许所有的VLAN通过，但是也不会有做为Super-VLAN的VLAN 10的数据帧从该接口进出。

（2）如果先配好了Trunk端口，并允许所有VLAN通过，则在此设备上将无法配置Super-VLAN。本质原因就是有物理端口的VLAN都不能被配置为Super-VLAN，而配置允许所有VLAN通过，则该Trunk端口就自动成为所有VLAN的成员，这样自然所有VLAN都不能配置为Super-VLAN了。对于图7-5中的Switch1而言，有效的VLAN只有VLAN 2和VLAN 3，所有的数据帧都在这两个VLAN中转发。

3．Sub-VLAN与外部网络的三层通信原理

前面说了，所有Sub-VLAN都是通过Super-VLAN的VLANIF接口作为网关与外部网络进行三层通信的。下面以图7-6所示的示例介绍Sub-VLAN与外部网络的三层通信原理。

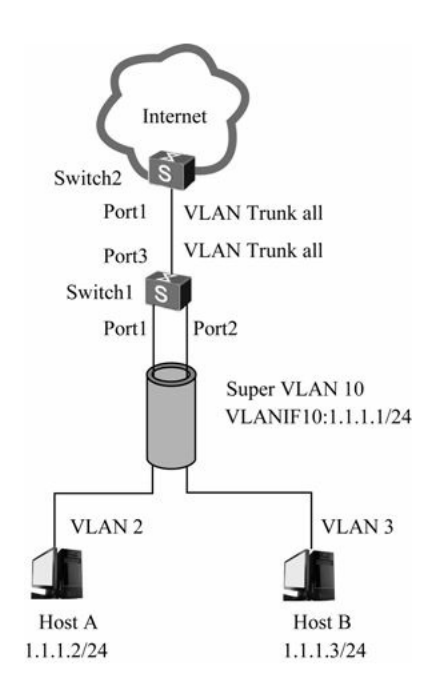


图7-5 Sub-VLAN与外部网络的二层通信示例

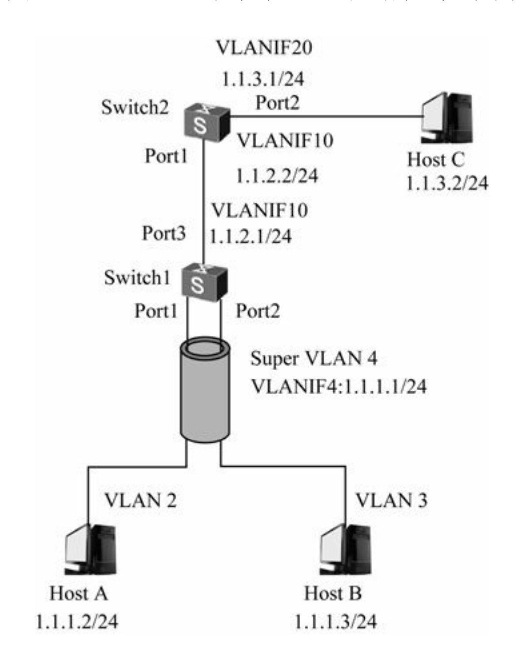


图7-6 Sub-VLAN与外部网络的三层通信示例

在本示例中，Switch1上配置了Super-VLAN 4，Sub-VLAN 2和Sub-VLAN 3，并配置一个普通的 VLAN 10；Switch2 上配置两个普通的 VLAN 10 和 VLAN 20。假设Super-VLAN 4中的Sub-VLAN 2下的主机A想访问与Switch2相连的主机C（位于VLAN 20中），则通信过程如下（假设Switch1上已配置了去往1.1.3.0/24网段的路由，Switch2上已配置了去往1.1.1.0/24网段的路由）。

（1）主机A将主机C的IP地址（1.1.3.2）和自己所在网段1.1.1.0/24进行比较，发现主机C和自己不在同一个子网。于是主机A发送一个ARP请求给自己的网关（位于Switch1上的Super-VLAN 4接口），请求网关的MAC地址。

（2）Switch1在收到该ARP请求后，查找Sub-VLAN和Super-VLAN的对应关系，从 Sub-VLAN 2 发送 ARP 应答给主机 A。ARP 应答数据帧中的源 MAC 地址为Super-VLAN 4对应的VLANIF4的MAC地址。

（3）主机A学习到网关的MAC地址后向网关发送目的MAC地址为Super-VLAN 4对应的VLANIF4的MAC地址、目的IP为1.1.3.2（主机C的IP地址）的数据帧。

（4）Switch1 在收到该数据帧后进行三层转发，根据在路由表中配置的路由表项知道其下一跳地址为1.1.2.2，出接口为VLANIF10，把数据帧发送给Switch2。

（5）Switch2在收到该数据帧后再进行三层转发，通过直连出接口VLANIF20，把数据帧发送给主机C。

（6）主机C在收到数据帧后向其网关（位于Switch2上的VLANIF20接口）发送应答数据帧，然后经过Switch2上进行三层转发到达Switch1。

（7）Switch1在收到来自主机C，并由Switch2转发的应答数据帧后再进行三层转发，通过其网关（VLANIF4接口）把数据帧发送给主机A。

以至完成Sub-VLAN与外部网络的三层通信全过程。

7.1.4 VLAN聚合配置思路

前面介绍了VLAN聚合的实现原理，Sub-VLAN间的三层通信原理，以及Sub-VLAN与外部网络间的二、三层通信原理。本节要介绍VLAN聚合的基本配置思路，也即基本配置任务。

从以上的介绍可以获知，要实现VLAN聚合功能，至少需要进行以下几方面的配置任务。

（1）创建Spuer-VLAN和各个Sub-VLAN。

（2）把各个 Sub-VLAN 中的用户计算机成员均配置在同一 IP 子网中，然后以基于端口划分方式加入对应的Sub-VLAN中。

（3）为Spuer-VLAN创建三层VLANIF接口，并且配置与各Sub-VLAN中用户计算机在同一IP子网中的IP地址。

（4）在 Spuer-VLAN 的三层 VLANIF 接口上使能 ARP 代理功能，以实现不同Sub-VLAN间的三层互通。

从以上配置任务可以看出，都是基于Spuer-VLAN和Sub-VLAN进行的配置，下面分别介绍。但一定要注意，必须先创建、配置各个 **Sub-VLAN**，再创建、配置**Spuer-VLAN**。

7.1.5 配置Sub-VLAN

在VLAN聚合中，Sub-VLAN的配置很简单，仅需要创建对应的Sub-VLAN，然后以基于端口划分方式把各用户计算机所连接的交换机端口加入对应的Sub-VLAN中。其实就是一个基于端口的VLAN划分配置过程。但要注意，各个Sub-VLAN的所有用户计算机网卡IP地址必须在同一个IP子网中。具体的配置步骤如表7-3所示。

表7-3 Sub-VLAN的配置步骤

步骤	命令	说明
1	**system-view** 例如：< HUAWEI > **system-view**	进入系统视图
2	**interface** *interface-type interface-number* 例如：[HUAWEI]**interface gigabitethernet 0/0/1**	键入要加入某 Sub-VLAN 的交换机端口，进入接口视图
3	**port link-type access** 例如：[HUAWEI -GigabitEthernet0/0/1] **port link-type access**	配置以上端口为 Access 类型。缺省为 Hybrid 类型，可用 **undo port link-type** 命令恢复为缺省的 Hybrid 类型
4	**quit** 例如：[HUAWEI -GigabitEthernet0/0/1] **quit**	返回系统视图
重复以上步骤，把其他要加入到 Sub-VLAN 的交换机端口转换成 Access 类型		
5	**vlan** *vlan-id* 例如：[HUAWEI] **vlan** 10	创建 Sub-VLAN，并进入 VLAN 视图。参数的取值范围是 1～4 094 的整数。缺省情况下，将所有端口都加入到一个缺省的 VLAN 1 中，可用 **undo vlan** *vlan-id* 删除指定 VLAN
6	**port** *interface-type* { *interface-number1*[**to** *interface-number2*] }&<1-10> 例如：[HUAWEI-VLAN10] **port** GigabitEthernet0/0/1**to** 0/0/4	将指定范围的交换机端口加入以上创建的 Sub-VLAN 中。但这里的交换机端口范围中各端口类型必须一样，如必须同为千兆以太网端口

7.1.6 配置Super-VLAN

Super-VLAN内可以包括多个Sub-VLAN，不能加入任何物理端口，但可以创建三层VLANIF接口并配置IP地址。另外，为了确保实现各Sub-VLAN间的三层通信，还需要在Super-VLAN的VLANIF接口上启用ARP代理功能。但要注意：在配置Super-VLAN之前必须已完成上节的Sub-VLAN配置。Super-VLAN的具体配置步骤如表7-4所示。

表7-4 Super-VLAN配置步骤

步骤	命令	说明
1	**system-view** 例如：< HUAWEI > **system-view**	进入系统视图
2	**vlan** *vlan-id* 例如：[HUAWEI] **vlan** 2	创建 Super-VLAN，并进入 VLAN 视图。本配置步骤中的 *vlan-id* 与 Sub-VLAN 中的 *vlan-id* 必须使用不同的 VLAN ID

（续表）

步骤	命令	说明
3	**aggregate-vlan** 例如：[HUAWEI -VLAN2] **aggregate-vlan**	将以上 VLAN 指定为 Super-VLAN。Super-VLAN 中不能包含任何物理端口，且 VLAN1 不能配置为 Super-VLAN 缺省情况下，没有配置当前 VLAN 为 Super-VLAN，可用 **undo aggregate-vlan** 命令恢复当前 VLAN 为普通 VLAN
4	**access-vlan** { *vlan-id1* [**to** *vlan-id2*] } &<1-10> 例如：[HUAWEI-vlan2] **access-vlan** 20 **to** 30	将上节配置好的 Sub-VLAN 加入第 2 步创建的 Super-VLAN 中。命令中的参数说明如下。 （1）*vlan-id1* 表示第一个 VLAN 的编号。取值范围是 1～4 094 的整数 （2）**to** *vlan-id2* 表示最后一个 VLAN 的编号。取值范围是 2～4 094 的整数，且 *vlan-id2* 的取值必须大于 *vlan-id1* 的取值，它和 *vlan-id1* 共同确定一个范围。如果将多个 Sub-VLAN 批量加入到 Super-VLAN 中，必须保证这些 Sub-VLAN 都没有创建对应的 VLANIF 接口 （3）&<1-10>：表示参数 *vlan-id1* [**to** *vlan-id2*]可以输入 1～10 次，但采用关键字 **to** 输入的不同 VLAN ID 区间必须没有交叉 【说明】**一个 VLAN 不能同时加入多个不同的 Super-VLAN 中**。不同的华为 S 系列交换机所支持聚合的 Sub-VLAN 数不一样，如 S2700 最多支持 16 个 Sub-VLAN 加入到同一 Super-VLAN 中，而 S7700&S9700 最多支持 256 个 Sub-VLAN 加入到同一 Super-VLAN 中。Super-VLAN 下 Sub-VLAN 过多，会存在 ARP 等广播复制性能问题，影响 ARP 学习，建议每个 Super-VLAN 下配置的 Sub-VLAN 数不超过 50 个 缺省情况下，Super-VLAN 中没有加入任何 Sub-VLAN，可用 **undo access-vlan** { *vlan-id1* [**to** *vlan-id2*] } &<1-10>命令将一个或一组 Sub-VLAN 从 Super-VLAN 中删除
5	**quit** 例如：[HUAWEI-vlan2] **quit**	退出 VLAN 视图，返回系统视图
6	**interface vlanif** *vlan-id* 例如：[HUAWEI] **interface vlanif** 2	键入 Super-VLAN 的 VLANIF 接口，进入接口视图
7	**arp-proxy inter-sub-vlan-proxy enable** 例如：[HUAWEI-Vlanif2] **arp-proxy inter-sub-vlan-proxy enable**	（可选）使能 Sub-VLAN 间的 ARP 代理功能。 如果需要在不同的 VLAN 间的实现三层互通，必须在接口上使能 VLAN 间的 ARP 代理功能 缺省情况下，关闭 VLAN 间 Proxy ARP 功能，可用 **undo arp-proxy inter-sub-vlan-proxy enable** 命令关闭 VLAN 间的 ARP 代理功能
8	**ip address** *ip-address* { *mask* | *mask-length* } [**sub**] 例如：[HUAWEI-Vlanif2] **ip address** 10.1.1.2 8	为以上 VLANIF2 接口配置主或从 IP 地址。命令中的参数说明如下。 （1）*ip-address*：指定 VLANIF 接口的 IP 地址； （2）*mask*：二选一参数，指定以上 VLANIF 接口 IP 地址对应的子网掩码； （3）*mask-length*：二选一参数，指定以上 VLANIF 接口 IP 地址对应的子网掩码长度 缺省情况下，在 VLANIF 接口上没有配置 IP 地址，可用 **undo ip address** *ip-address* { *mask* | *mask-length* } [**sub**] 命令删除指定的 IP 地址

【示例 】 向Super-VLAN 2中加入Sub-VLAN 20和Sub-VLAN 30。

<HUAWEI> system-view

[HUAWEI] vlan 2

[HUAWEI-vlan2] aggregate-vlan

[HUAWEI-vlan2] access-vlan 20 30

【示例 2】使能VLANIF接口上的VLAN间ARP代理功能。

<HUAWEI>system-view

[HUAWEI] interface vlanif 10

[HUAWEI-Vlanif10] arp-proxy inter-sub-vlan-proxy enable

7.1.7 VLAN聚合配置示例

如图 7-7 所示，某公司拥有多个部门且位于同一IP子网。为了提升业务的安全性，已将不同部门的用户划分到不同VLAN中。现由于业务需要，不同部门（如VLAN 2和VLAN 3为不同部门）间的用户需要实现

三层互通。

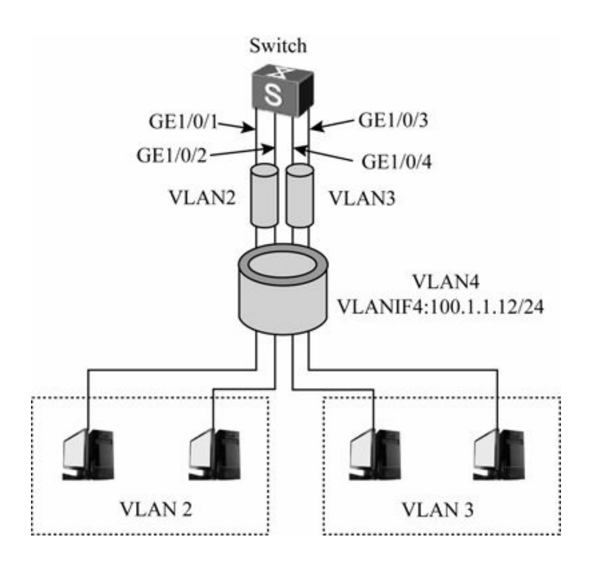


图7-7 VLAN聚合示例

本示例的配置很简单，可以在Switch上部署VLAN聚合，同时在Super-VLAN的VLANIF接口上启用ARP代理功能，即可在满足VLAN 2和VLAN 3二层隔离的同时实现三层互通，而且还满足了VLAN 2和VLAN 3中的用户计算机 IP地址配置在同一个 IP 子网之中，节省了IP地址。

下面是具体的配置步骤。

（1）转换连接Sub-VLAN 2和Sub-VLAN 3用户的交换机端口类型为Access类型，因为华为S系列交换机的以太网接口缺省均为Hybrid类型。

```
<HUAWEI>system-view
[HUAWEI] interface gigabitethernet 1/0/1
[HUAWEI-GigabitEthernet1/0/1] port link-type access
[HUAWEI-GigabitEthernet1/0/1] quit
[HUAWEI] interface gigabitethernet 1/0/2
[HUAWEI-GigabitEthernet1/0/2] port link-type access
[HUAWEI-GigabitEthernet1/0/2] quit
[HUAWEI] interface gigabitethernet 1/0/3
[HUAWEI-GigabitEthernet1/0/3] port link-type access
[HUAWEI-GigabitEthernet1/0/3] quit
[HUAWEI] interface gigabitethernet 1/0/4
[HUAWEI-GigabitEthernet1/0/4] port link-type access
[HUAWEI-GigabitEthernet1/0/4] quit
```

（2）创建Sub-VLAN 2，并向其中加入GE1/0/1和GE1/0/2端口。

```
[HUAWEI] vlan 2
[HUAWEI-vlan2] port gigabitethernet 1/0/1 1/0/2
[HUAWEI-vlan2] quit
```

（3）创建Sub-VLAN 3，并向其中加入GE1/0/3和GE1/0/4端口。

[HUAWEI] vlan 3

[HUAWEI-vlan3] port gigabitethernet1/0/3 1/0/4

[HUAWEI-vlan3] quit

（4）创建Spuer-VLAN 4，并聚合VLAN 2和VLAN 3。

[HUAWEI] vlan 4

[HUAWEI-vlan4] aggregate-vlan

[HUAWEI-vlan4] access-vlan 2 to 3

[HUAWEI-vlan4] quit

（5）配置Spuer-VLAN 4的VLANIF4接口 IP地址，并启用ARP代理功能。

[HUAWEI] interface vlanif4

[HUAWEI-Vlanif4] ip address 100.1.1.12 255.255.255.0

[HUAWEI-Vlanif4] arp-proxy inter-sub-vlan-proxy enable

[HUAWEI-Vlanif4] quit

配置完成后，VLAN2的用户与VLAN3的用户可以相互ping通。

【经验之谈】 本示例仅是一个最基本的 VLAN 聚合示例，因为它是在一台三层交换机上实现的。事实上 VLAN 聚合也可以实现跨交换机的 VLAN 聚合。如图 7-8 所示， SwitchA是一台汇聚层的三层交换机，创建一个 Super-VLAN 10，并配置其VLANIF10接口 IP地址为 10.1.1.1/24，同时创建作为Sub-VLAN的VLAN 2和VLAN 3；SwitchB、SwitchC 和 SwitchD 是三层接入层交换机，可以是二层或三层交换机，分别创建了以上对应的VLAN 2、VLAN 3，且各Sub-VLAN中的用户计算机 IP地址都位于 10.1.1.0/24的IP子网中。在SwitchB、SwitchC和SwitchD上把与PC连接的端口配置为Access类型，把与SwitchA连接的端口配置为Trunk类型，并允许所连接的Sub-VLAN通过；把SwitchA与SwitchB、SwitchC和SwitchD连接的端口配置为Trunk类型，也允许所连接的Sub-VLAN通过。

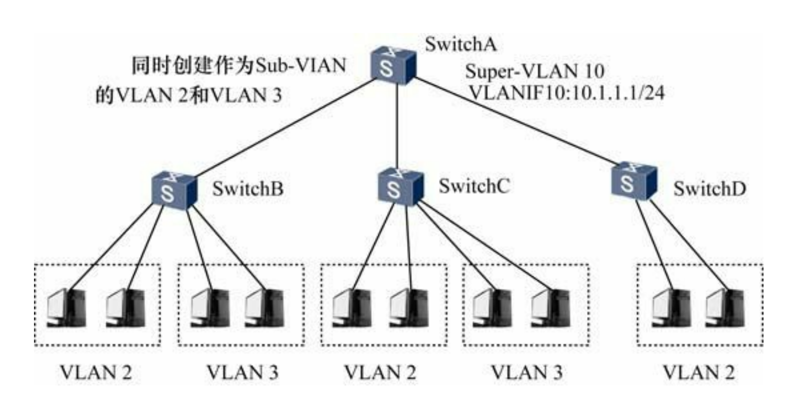


图7-8 跨交换机的VLAN聚合示例

说明

在VLAN聚合配置管理中，除了可以在任意视图下使用我们在本书第6章已 经介绍的 display vlan [vlan-id [verbose]] 命令查看所有VLAN或指定VLAN的相关信息，使用 display interface vlanif [vlan-id] 命令查看Super-VLAN的VLANIF接口信息外，还可使用 display sub-vlan [vlan-id] 命令查看 Sub-VLAN类型的VLAN表项信息，使用 display super-vlan [vlan-id] 命令查看 Super-VLAN类型的VLAN表项信息。

7.2 MUX VLAN配置与管理

在 VLAN 的应用过程中经常遇到这样的需求：整个公司网络的用户都处于一个 IP子网中，但又希望在所有员工都能直接二层访问网络中的某关键设备的同时一部分员工彼此之间二层隔离。例如，在企业网络中，企业员工和企业客户可以访问企业的服务器，但是仅希望企业内部员工之间可以互相交流，而企业客户之间是隔离的。这时如果仅仅考虑到普通 VLAN 就很难实现了，因为如果企业规模很大，拥有大量的用户，那么就要为不能互相访问的用户都分配VLAN（特别是在 ISP中），这不仅需要耗费大量的VLAN ID，还在增加了网络管理者的工作量同时也增加了维护量。通过本节将要介绍的MUX VLAN （Multiplex VLAN，复合VLAN）提供的二层流量隔离机制就可以实现以上双重目的。

7.2.1 MUX VLAN概述

MUX VLAN提供了一种通过 VLAN进行网络资源访问控制的机制。它与上节介绍的VLAN聚合在形式上有些类似，也包括两个层次的VLAN，即Principal VLAN（主VLAN）和Subordinate VLAN（从VLAN），但也只是一种关联关系，不是内层和外层VLAN关系。从VLAN又分为两类，即Separate VLAN（隔离型从VLAN）和Group VLAN（互通型从VLAN）。

本项VLAN特性除了S1700、S2700系列外的其他所有华为S系列交换机均支持。

MUX VLAN具有以下特性。

（1）主VLAN可以与任何从VLAN间直接二层通信。

（2）任何不同从VLAN（包括隔离型从VLAN和互通型从VLAN）间都不能直接二层通信。

（3）互通型从 VLAN 内部的用户间可直接二层通信，隔离型从 VLAN 内部的用户间不能直接通信，起到在同一个VLAN内实现各用户相互二层隔离的目的。

以上这几种不同类型的MUX VLAN的基本特性如表 7-5所示。

表7-5 MUX VLAN中的不同VLAN类型及基本特性

MUX VLAN	VLAN 类型	所属端口	通信权限
Principal VLAN（主 VLAN）	—	Principal port（主端口）	Principal port 可以和 MUX VLAN 内的所有端口直接进行二层通信
Subordinate VLAN（从 VLAN）	Separate VLAN（隔离型从 VLAN）	Separate port（隔离型从端口）	Separate port 只能和 Principal port 进行二层通信，和其他类型的端口实现完全隔离。这是为了实现单个用户间隔离的目的而推出的 VLAN 技术，每个 Separate VLAN 必须绑定一个 Principal VLAN
	Group VLAN（互通型从 VLAN）	Group port（互通型从端口）	Group port 可以和 Principal port 进行通信，在同一互通型从 VLAN 内的端口也可相互进行二层通信，但不能和其他互通型从 VLAN 的端口或 Separate port 通信。这是为了实现不同用户组间隔离，同一个用户组内部互通的目的而推出的 VLAN 技术，每个 Group VLAN 必须绑定一个 Principal VLAN

图 7-9是MUX VLAN的一种典型应用，企业可以用Principal port连接用户需要共同访问的企业服务器，Separate port连接企业客户，Group port连接企业员工。这样就能够实现企业客户、企业员工都能够访问企业服务器，而企业员工内部可以通信、企业客户间不能通信、企业客户和企业员工之间不能互访的目的。

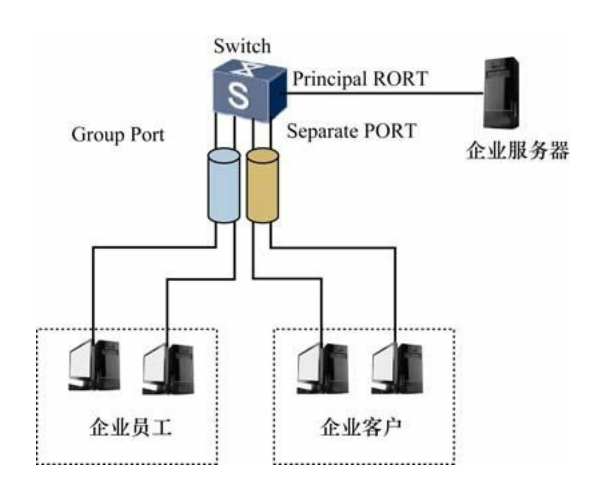


图7-9 MUX VLAN典型应用示例

7.2.2 配置MUX VLAN

MUX VLAN的配置也很简单，基本配置任务就是以下两项。

（1）创建一个主VLAN，一个或多个隔离型从VLAN和互通型VLAN;

（2）在加入以上各VLAN的交换机端口上使能MUX VLAN功能。

1．配置MUX VLAN中的主VLAN

MUX VLAN中主VLAN（Principal VLAN）的配置方法很简单，具体步骤如表 7-6所示。

表7-6 主VLAN的配置步骤

步骤	命令	说明
1	**system-view** 例如：< HUAWEI > **system-view**	进入系统视图
2	**vlan** *vlan-id* 例如：[HUAWEI] **vlan** 2	创建主 VLAN，并进入 VLAN 视图
3	**mux-vlan** 例如：[HUAWEI-VLAN2] **mux-vlan**	将以上 VLAN 指定为 MUX VLAN 的主 VLAN。缺省情况下，没有配置当前 VLAN 为主 VLAN，可用 **undo mux-vlan** 命令取消当前 VLAN 为主 VLAN

注意

当指定VLAN ID已经用于主VLAN，则该VLAN不能配置VLANIF接口，也不 能配置VLAN Mapping、VLAN Stacking、Super-VLAN、Sub-VLAN功能。如果某VLAN的类型已配置为Super VLAN、Sub VLAN或从VLAN，则该VLAN不能配置为主VLAN；如果某VLAN上已创建了VLANIF接口，则该VLAN也不能配置为主VLAN。

【示例 1】将VLAN5配置为MUX VLAN中的主VLAN。

<HUAWEI>system-view

[HUAWEI] vlan 5

[HUAWEI-vlan5] mux-vlan

2．配置MUX VLAN中的从VLAN

前面已介绍了，从VLAN又分为互通型从VLAN（Group VLAN）和隔离型从VLAN （Separate VLAN）两类。但一个MUX VLAN不一定要求同时包括这两种从VLAN，且一个主**VLAN**下只能配置一个隔离型从**VLAN**，却可以最多配置128个互通型从VLAN。

互通型从VLAN可实现同一VLAN内用户端口间的相互通信；隔离型从VLAN可隔离同一VLAN内用户端口间的相互通信。但同一MUX VLAN中互通型从VLAN和隔离型从VLAN的VLAN ID不能一样。

从VLAN的配置方法也很简单，具体如表7-7所示。

表7-7 从VLAN的配置步骤

步骤	命令	说明
1	**system-view** 例如：< HUAWEI > **system-view**	进入系统视图
2	**vlan** *vlan-id* 例如：[HUAWEI] **vlan** 2	创建各个从 VLAN，并进入 VLAN 视图
3	**quit** 例如：[HUAWEI-VLAN2] **quit**	返回系统视图
4	**vlan** *vlan-id* 例如：[HUAWEI] **vlan** 5	进入主 VLAN 视图
5	**subordinate group** { *vlan-id1* [**to** *vlan-id2*] } &<1-10> 例如：[HUAWEI-VLAN2] **subordinate group** 2 3	把指定范围的 VLAN 配置为互通型从 VLAN。缺省情况下，没有配置主 VLAN 下的互通型从 VLAN，可用 **undo subordinate group** 命令删除主 VLAN 下的互通型从 VLAN
6	**subordinate separate** *vlan-id* 例如：[HUAWEI-vlan2] **subordinate separate** 4	将指定 VLAN 配置为隔离型从 VLAN。缺省情况下，没有配置主VLAN下的隔离型从VLAN，可用 **undo subordinate separate** 命令删除主 VLAN 下的隔离型从 VLAN

注意

当指定VLAN ID已经用于互通型从VLAN或隔离型从VLAN，或者主VLAN， 则该 VLAN不能配置VLANIF接口，也不能配置 VLAN Mapping、VLAN Stacking、Super-VLAN、Sub-VLAN功能。如果某VLAN的类型已配置为Super VLAN、Sub VLAN或从VLAN，则该VLAN不能配置为互通型从VLAN和隔离型从VLAN；如果某VLAN上已创建了VLANIF接口，则该VLAN也不能配置为互通型从VLAN和隔离型从VLAN。

【示例 2】将VLAN 7～10配置为VLAN 5的互通型从VLAN。

<HUAWEI>system-view

[HUAWEI] vlan 5

[HUAWEI-vlan5] subordinate group 7 to 10

【示例 3】将VLAN 6配置为VLAN 5的隔离型从VLAN。

<HUAWEI>system-view

[HUAWEI] vlan 5

[HUAWEI-vlan5] subordinate separate6

3．使能端口MUX VLAN功能

只有使能端口MUX VLAN功能后，才能达到主VLAN与从VLAN之间通信、互通型从 VLAN 内的端口可以相互通信和隔离型从 VLAN 端口间不能相互通信的目的。但要特别注意的是，在MUX VLAN中的所有终端设备连接的交换机端口必须是以**Access**类型或**Hybrid Untagged**类型 加入的，且端口只能允许一个**VLAN**通过 。如果端口允许多个VLAN通过，那么端口将无法成功使能MUX VLAN功能。

在交换机端口上使能MUX VLAN功能的配置很简单，就是在对应的交换机端口视图下执行**port mux-vlan enable** 命令（注意要在所有**MUX VLAN**的交换机端口上配置 ）。端口使能MUX VLAN功能后，该端口不可以再用于VLAN Mapping、VLAN Stacking配置。

注意

禁止端口MAC地址学习功能或限制端口MAC地址学习数量会影响MUX VLAN功能的正常使用。所以，不能在同一端口上同时配置MUX VLAN和接口安全功能；不能在同一端口上同时配置MUX VLAN和MAC认证功能；不能在同一端口上同时配置MUX VLAN和 802.1x认证功能。

【示例 4】使能GE0/0/1端口上的MUX VLAN功能。

```
<HUAWEI>system-view
[HUAWEI] interface gigabitethernet 0/0/1
[HUAWEI-GigabitEthernet0/0/1] port mux-vlan enable
```

【说明】在MUX VLAN配置管理中，除了可以在任意视图下使用本书第 6章已经介绍的display vlan [vlan-id [verbose]] 命令查看所有VLAN或指定VLAN的相关信息，使用display interface vlanif [vlan-id] 命令查看Super-VLAN的VLANIF接口信息外，还可使用display mux-vlan命令查看所有MUX VLAN的相关信息。可查的MUX VLAN配置信息包括主VLAN ID、从VLAN ID、VLAN的类型、VLAN包含的交换机端口。

7.2.3 MUX VLAN配置示例

本示例拓扑结构如图7-10所示。在某小型企业网络中，要求通过 MUX VLAN功能实现企业所有员工都可以访问企业的服务器（Server），但希望企业内部部分员工之间可以互相通信，而另一部分员工之间是隔离的，不能够互相访问。

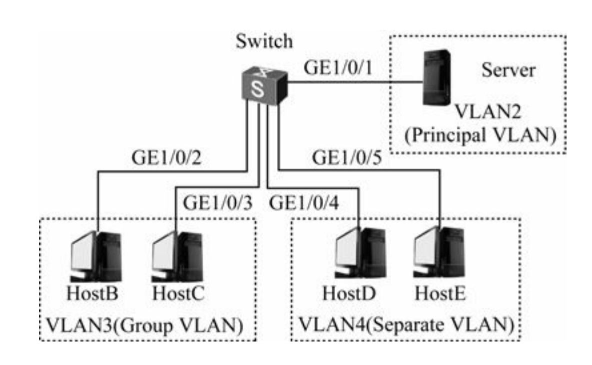


图7-10 MUX VLAN配置示例拓扑结构

因为MUX VLAN的配置思路很简单，各配置任务中的配置步骤也很简单，所以在此直接介绍本示例的配置步骤。

（1）创建各个MUX VLAN，即VLAN2、VLAN3和VLAN4。

```
<HUAWEI>system-view
[HUAWEI] vlan batch 2 to 4
```

（2）配置VLAN 2为主VLAN，VLAN 3为互通型从VLAN，VLAN 4为隔离型从VLAN。

```
[HUAWEI] vlan 2
[HUAWEI-vlan2] mux-vlan
[HUAWEI-vlan2] subordinate group 3
[HUAWEI-vlan2] subordinate separate4
[HUAWEI-vlan2] quit
```

（3）配置各MUX VLAN中的交换机端口为Access类型（也可以是仅允许一个VLAN通过的Untagged Hybrid类型），加入对应的VLAN中，并使能它们的MUX VLAN功能。

```
[HUAWEI] interface gigabitethernet 1/0/1
```

```
[HUAWEI-GigabitEthernet1/0/1] port link-type access
[HUAWEI-GigabitEthernet1/0/1] port default vlan 2
[HUAWEI-GigabitEthernet1/0/1] port mux-vlan enable
[HUAWEI-GigabitEthernet1/0/1] quit
[HUAWEI] interfacegigabitethernet 1/0/2
[HUAWEI-GigabitEthernet1/0/2] port link-type access
[HUAWEI-GigabitEthernet1/0/2] port default vlan 3
[HUAWEI-GigabitEthernet1/0/2] port mux-vlan enable
[HUAWEI-GigabitEthernet1/0/2] quit
[HUAWEI] interface gigabitethernet 1/0/3
[HUAWEI-GigabitEthernet1/0/3] port link-type access
[HUAWEI-GigabitEthernet1/0/3] port default vlan 3
[HUAWEI-GigabitEthernet1/0/3] port mux-vlan enable
[HUAWEI-GigabitEthernet1/0/3] quit
[HUAWEI] interfacegigabitethernet 1/0/4
[HUAWEI-GigabitEthernet1/0/4] port link-type access
[HUAWEI-GigabitEthernet1/0/4] port default vlan 4
[HUAWEI-GigabitEthernet1/0/4] port mux-vlan enable
[HUAWEI-GigabitEthernet1/0/4] quit
[HUAWEI] interface gigabitethernet 1/0/5
[HUAWEI-GigabitEthernet1/0/5] port link-type access
[HUAWEI-GigabitEthernet1/0/5] port default vlan 4
[HUAWEI-GigabitEthernet1/0/5] port mux-vlan enable
[HUAWEI-GigabitEthernet1/0/5] quit
```

以上配置完成后，可以通过简单的Ping操作，验证Server和HostB、HostC、HostD、HostE都可以互相ping通；HostB和HostC可以互相ping通；HostD和HostE不可以互相ping通；HostB、HostC和HostD、HostE不可以互相ping通。符合在7.2.1节介绍的MUX VLAN主要特性。

说明

本示例仅是最简单的MUX VLAN应用示例，其实MUX VLAN同样可应用于如图 7-8所示的跨交换机的VLAN环境中，只是需要注意在各MUX VLAN中的端口成员必须是Access类型，但交换机之间连接的端口仍然是Trunk或者Hybrid类型的，且可以允许多个VLAN通过。

7.3 QinQ基础

我们在此之前介绍的VLAN都是单层标签的，也就是在数据帧中只有一个802.1Q标签头，本节介绍的QinQ（是802.1Q-in-802.1Q的简称）技术是一项可在数据帧中的原802.1Q标签头基础上再增加一层802.1Q标签头，实现双VLAN标签的目的。但要注意，启用了 QinQ功能的交换机端口具有添加和剥离外层**VLAN** 标签的双重功能，即上行传输时对帧添加外层标签，而下行传输时又可以剥离帧中的外层标签 ，以实现正常的流量转发。那么这样的双VLAN标签有什么意义呢？这就需要从QinQ技术产生的背景来进行分析了。

7.3.1 QinQ技术诞生的背景

QinQ最初主要是为扩展VLAN ID空间而产生的，但随着城域以太网的发展以及运营商精细化运作的要求，QinQ的双层标签有了进一步的使用场景。它的内、外层标签可以代表不同的信息，如内层标签代表用户，外层标签代表业务。另外，QinQ数据帧带着两层标签穿越运营商网络，内层标签透明传送，也可以看作是一种简单、实用的VPN技术。因此，它又可以作为核心MPLS VPN在城域以太网VPN的延伸，最终形成端到端的VPN技术。由于QinQ方便易用的特点，现在已经在各运营商中得到了广泛的应用，如QinQ技术在城域以太网解决方案中和多种业务相结合。特别是灵活QinQ（Selective QinQ/VLAN Stacking）的出现，使得QinQ业务更加受到了运营商的推崇和青睐。

我们知道，普通 VLAN 中的一个 VLAN 标签是用来区分用户的，但如果想要同时区分用户和业务类型，那么怎么办呢？图 7-11 是一个总公司下面连接了两个分支子公司，而各分支子公司中已对不同部门的员工采用了VLAN进行区分，但两个子公司的部门VLAN ID规划是重叠的。这样如果数据帧中只采用一层VLAN标签，总公司就无法区分数据是来自哪个子公司的也就无法针对不同子公司的数据进行任何处理了。

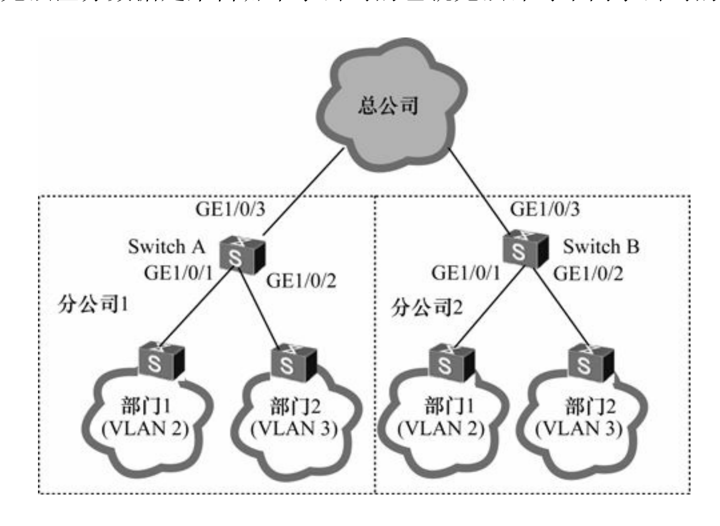


图7-11 QinQ典型应用示例

为了解决这个问题，可以设想在总公司的交换机上为各子公司创建了不同的VLAN。这样当连接对应子公司的总公司交换机端口收到数据帧后再在数据帧外面添加一层VLAN标签（此时数据帧中就有两层VLAN标签了，原来的VLAN标签称之为内层VLAN标签，新添加的称之为外层VLAN标签），如果为子公司1和子公司2的数据帧分别添加的外层VLAN标签为VLAN 10和VLAN 20，就可实现在总公司中对来自不同子公司的数据进行区分了，也可以对来自这两个子公司的数据提供不同的服务，即差分服务了。

另外，在基于传统的 802.1Q 协议的二层局域网互联模式中，当两个用户网络需要通过服务提供商（ISP）互相访问时（如在城域以太网中），ISP 必须为每个接入用户创建不 同的 VLAN。这种配置方法一方面使得用户的 VLAN 在骨干网络上可见，存在一定的安全隐患，同时因为一一对应的VLAN ID，也消耗了大量服务提供商的VLAN ID资源。这对较大的 ISP来说是无法承受的，因为只有4 094个VLAN ID可用），当接入的用户数目很多时可能使 ISP网络的VLAN ID不够用。另外，采用这种普通VLAN部署方式下，不同的 ISP接入用户就不能使用相同的VLAN ID，否则就无法实现不同接入用户间的隔离，这时用户的VLAN ID只能由 ISP统一规划，导致用户没有自己规划VLAN的权利。

通过QinQ技术可以有效地解决以上问题，因为它可以为许多不同内层VLAN标签用户使用同一个外层VLAN标签进行封装，解决了 ISP的VLAN ID资源不足的问题。另外，通过外层VLAN标签对内层VLAN标签的屏蔽作用，使用户自己的内层VLAN ID部署可以由用户自己作主，而不必由ISP来统一部署。

这个双层 VLAN 标签可以当作单层 VLAN 标签使用，即仅使用新添加的外层公网VLAN标签，内层私网VLAN可以作为数据来传输，如在本章后面将要介绍的 2 to 1的VLAN映射中；当然也可以作为双层VLAN标签来使用（如在本章后面将要介绍的 2 to 2的VLAN映射中），整个数据帧中的VLAN标签由内、外双层VLAN标签共同决定，这样一来，就相当于可以使用的VLAN ID数量达到了 4 094×4 094个了，以此来达到扩展VLAN空间的目的。通过这样的双层VLAN标签封装，可以使私网VLAN ID在公网上透传，既解决了用户VLAN ID的安全性问题和由用户自己规划私网VLAN ID的需求问题，又解决了 ISP的VLAN ID空间不足的问题。因为在 ISP中可以为需要相互访问的用户配置相同的外层VLAN，也只需为来自同一用户网络的不同VLAN提供一个VLAN ID。

7.3.2 QinQ封装和终结

QinQ是在传统802.1Q VLAN标签头基础上再增加一层新的802.1Q VLAN标签头，如图7-12所示。由此可知，QinQ帧比传统的802.1Q帧多了4个字节，即新增的802.1Q VLAN标签。

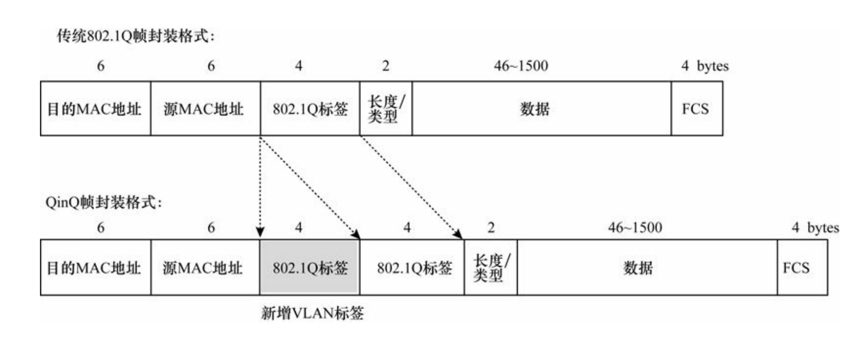


图7-12 传统802.1Q帧和QinQ帧格式比较

QinQ帧封装的过程就是把单层802.1Q标签的数据帧转换成双层802.1Q标签的数据帧。封装过程主要发生在城域网侧连接用户的交换机端口上。根据不同的VLAN标签封装依据，QinQ可以分为“基本QinQ”和“灵活QinQ”两种类型。具体说明如下。

1．基本QinQ封装

“基本QinQ封装”是将进入一个端口的所有流量全部封装一个相同的外层VLAN标 签，是一种基于端口的QinQ封装方式，也称“QinQ二层隧道”。开启端口的基本QinQ功能后，当该端口接收到已经带有VLAN标签的数据帧时，则该数据帧就将封装成双层标签的帧；如果接收到的是不带VLAN标签的数据帧，则该数据帧将封装成为带有端口缺省VLAN的一层标签的帧。

从以上介绍可以看出，基本QinQ的VLAN标签封装不够灵活，很难有效地区分不同的用户业务，因为它对进入同一个交换机端口的所有数据帧都封装相同的外层VLAN标签。但在需要较多的VLAN时，可以使用这个基本QinQ功能，这样可以减少对VLAN ID的需求，因为进入同一个端口的所有数据帧都封装同一个外层VLAN标签。

如图7-13所示的网络中，企业部门1（Department1）有两个办公地，部门2（Department2）有三个办公地，两个部门的各办公地分别和网络中的PE1、PE2相连，部门1和部门2可以任意规划自己的VLAN。这样，可在PE1和PE2上通过如下思路配置QinQ二层隧道功能，使得每个部门的各个办公地网络可以互通，但两个部门之间不能互通。

（1）在PE1上，对于进入端口Port1和Port2的用户（都属于部门1）数据帧都封装外层VLAN 10，对于进入端口Port3中的用户（属于部门 2）数据帧都封装外层VLAN 20。

（2）在PE2上，对于进入端口Port1和Port2的用户（都属于部门2）数据帧都封装外层VLAN 20。

（3）PE1上的端口Port4和PE2上的端口Port3允许VLAN 20的用户数据帧通过，以便实现连接在PE1的Port3上部门2的用户与连接在PE2的Port1和Port2上部门2的用户互通。

这种基本 QinQ 封装就相当于用一个外层的 VLAN 标签映射同类用户的多个内层VLAN标签，以减少ISP端设备VLAN ID的使用量。

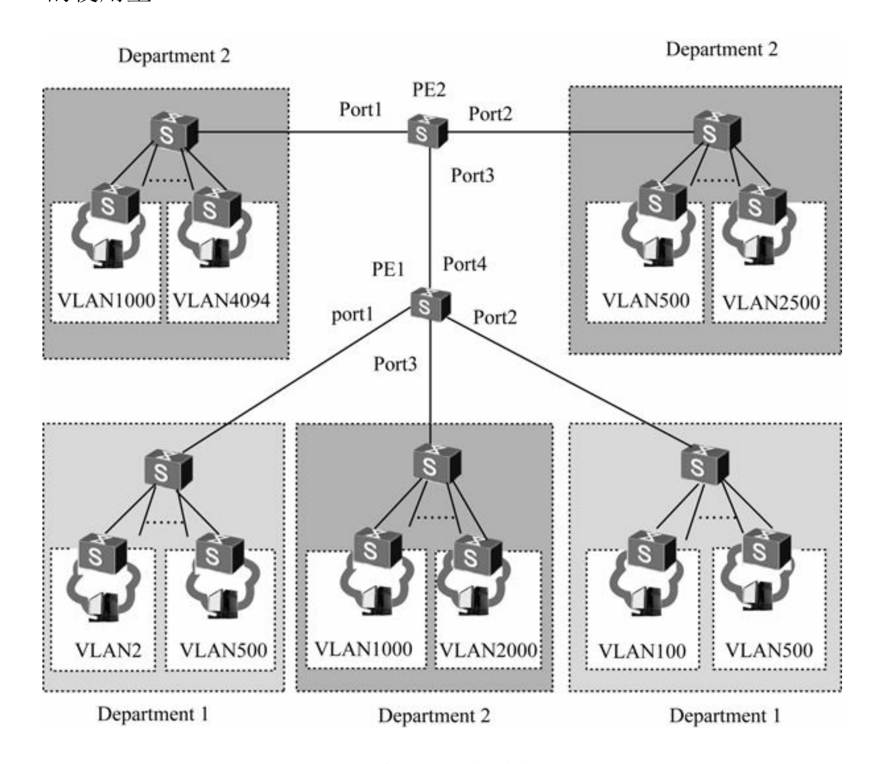


图7-13 基本QinQ典型应用示例

2．灵活QinQ封装

“灵活QinQ”是对QinQ的一种更灵活的实现，是基于端口封装与基于VLAN封装的结合方式。除了能实现所有基本QinQ的功能外，灵活QinQ对于同一个端口接收的数据帧还可以根据不同的内层VLAN标签执行不同的外层标签封装。它又可分为以下3个子类。

（1）基于VLAN ID的灵活QinQ：它是基于数据帧中不同的内层标签的VLAN ID来添加不同的外层标签。即具有相同内层标签的帧添加相同的外层VLAN标签，具有不同内层标签的帧添加不同的外层 VLAN标签。这就要求不同用户的内层 **VLAN ID**或**VLAN ID**范围绝对不能重叠或交叉 。华为S系列交换机中的**S2700**、**S3700**、**S5700**、**S6700**仅支持基于**VLAN ID**的灵活**QinQ**功能 。

（2）基于802.1p优先级的灵活QinQ：它是基于数据帧中不同的内层标签的802.1p优先级来添加不同的外层标签。即具有相同内层VLAN 802.1p优先级的帧添加相同的外层标签，具有不同内层VLAN 802.1p优先级的帧添加不同的外层标签。这就要求不同用户的内层**VLAN**的**802.1p**优先级或**802.1p**优先级范围绝对不能重叠或交叉 。基于802.1p优先级的灵活QinQ在华为S系列交换机中仅**S7700**、**S9300**和**S9700**系列支持 。

（3）基于流策略的灵活QinQ：它是根据所定义的QoS策略为不同的数据帧添加不同的外层标签。基于流策略的灵活QinQ是基于端口与VLAN相结合的方式实现的，能够针对业务类型提供差别服务。基于流策略的灵活QinQ在华为S系列交换机中仅**S7700**、**S9300**和**S9700**系列支持 。

以上三种灵活QinQ的配置方法将在本章后面具体介绍。

当同一用户的不同业务需要使用不同的VLAN ID时，可以根据VLAN ID区间进行分流。现假设PC上网的VLAN ID范围是101～200；IPTV的VLAN ID范围是201～300；大客户的VLAN ID范围是301～400。面向用户的端口在收到用户数据后根据用户VLAN ID范围，对PC上网业务封装上外层标签100，对IPTV封装上外层标签300，对大客户封装上外层标签500。

说明

QinQ封装一般在交换式端口上进行，但也可以在路由子接口上进行（QinQ终结只能在路由子接口上进行）。此种方法可以通过一个子接口来透传多个标识用户的VLAN ID，这种子接口也叫QinQ Stacking子接口。这种封装方式也是基于流的QinQ封装方式，但QinQ Stacking子接口只能和L2VPN业务结合起来才有意义，不支持三层转发功能。

在如图7-14所示的网络中，企业的部门1有多个办公地，部门2也有多个办公地。部门 1的网络中使用VLAN 2～VLAN 500；部门 2的网络中使用VLAN 500～VLAN 4094。PE1的Port1端口会同时收到两个部门不同VLAN区间的用户数据帧。

此时可根据图中标识的各办公地的用户VLAN ID范围在PE1和PE2上通过如下思路配置基于VLAN的灵活QinQ功能，使得每个部门的各个办公地网络之间可以互通，但两个部门之间不能互通。具体配置思路如下。

（1）对于进入PE1的Port1端口的用户数据帧，依据其VLAN ID的不同添加对应的外层VLAN标签。如VLAN ID在 2～500之间，则封装VLAN ID为 10的外层标签；如VLAN ID在 1 000～2 000之间，则封装VLAN ID为 20的外层标签。

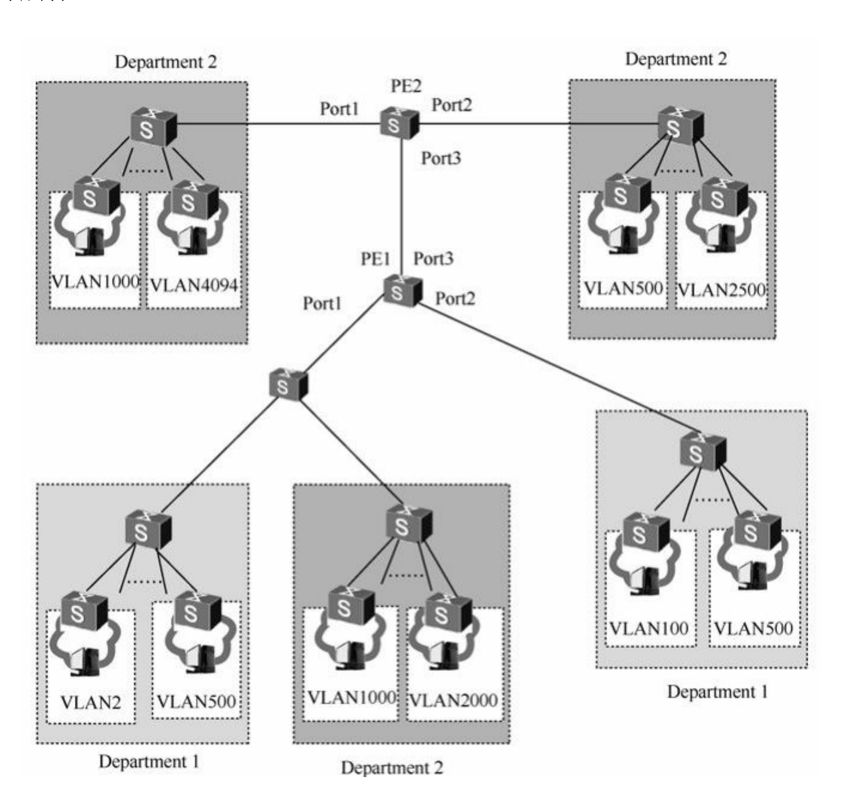


图7-14 灵活QinQ典型应用示例

（2）对于进入PE1的Port2端口的用户数据帧，如果VLAN ID在 100～500之间，则封装VLAN ID为 10的外层标签。

（3）对于进入PE2的Port1端口的用户数据帧，如VLAN ID在 1 000～4 094之间，则封装VLAN ID为 20的外层标签。

（4）对于进入PE2的Port2端口的用户数据帧，如果VLAN ID在 500～2 500之间，则封装VLAN ID为 20的外层标签。

（5）在 PE1和 PE2的 Port3端口上允许VLAN 20的帧通过，以便实现连接在 PE1的Port1端口下连接的部门2用户与连接在PE2的Port1和Port2的部门2的用户互通。

从以上可以看出，灵活QinQ比基本QinQ的外层标签封装更加灵活，可以根据用户数据帧中原来的VLAN ID范围来确定封装不同的外层标签，这样更方便了对相同网络中不同业务的用户数据流提供差分服务。

3．QinQ/Dot1q终结子接口

QinQ/Dot1q 终结是指设备对数据帧的双层或者单层 VLAN 标签进行识别，根据后续的转发行为对帧中的双层或者单层 VLAN 标签进行剥离，然后继续传送。也就是这些 VLAN 标签仅在此之前生效，后面的数据传输和处理不再依据帧中的这些VLAN标签。

终结一般在路由子接口上执行，即终结子接口，如我们在单臂路由中就要配置路由子接口的802.1Q的VLAN终结。如果路由子接口是对数据帧的单层VLAN标签终结， 那么该子接口称为 Dot1q 终结子接口；如果路由子接口是对数据帧的双层 VLAN 标签终结，那么该子接口称为QinQ终结子接口。QinQ终结子接口根据终结的用户VLAN标签的类型，通常分为两种子接口。

（1）明确的QinQ终结子接口：两层VLAN标签为固定的值。

（2）模糊的QinQ终结子接口：两层VLAN标签为范围值，即终结的内、外层标签都为一个VLAN ID范围值。

7.3.3 TPID的可调值

TPID（Tag Protocol Identifier，标签协议标识）是图 7-12所示的VLAN帧中的 802.1Q标签头中的一个字段，表示 VLAN 标签的协议类型。该字段的缺省值为 0x8100，用来标识对应帧是一个带有 IEEE 802.1Q标签的帧，如图 7-15所示。但某些厂商将设备可识别的TPDI字段值设为0x9100或其他值。

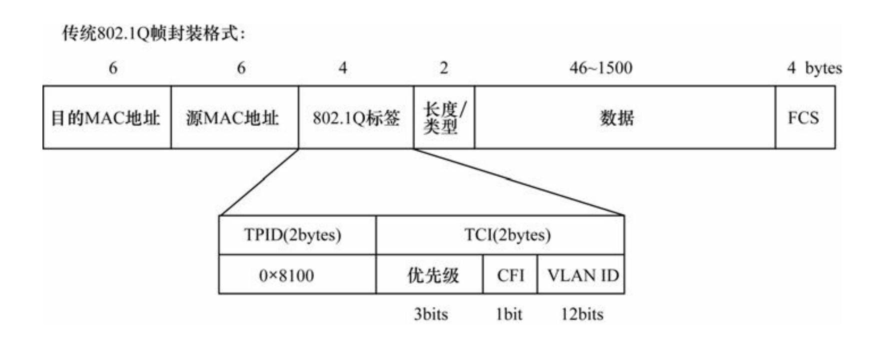


图7-15 VLAN帧中的TPID字段

802.1Q标签头位于“源MAC地址”和“长度/类型”（Length/Type）字段之间。通过检查外层标签中的TPID字段值，设备可确定收到的帧承载的是运营商VLAN标签，还是用户VLAN标签。设备在接收到帧后，将设备自身配置的TPID值与帧中最外层的VLAN标签的TPID字段值进行比较。如果能与帧中的TPID字段的值匹配，则该帧承载的是对应的VLAN标签。

例如，如果帧中仅一层标签匹配的话，则肯定是用户VLAN标签；如果是双层标签的QinQ帧，则仅与

新添加的外层标签中的TPID字段值进行比较，如果一致则证明帧中承载的是运营商的VLAN标签。如一个数据帧中承载的外层和内层标签中的TPID字段值分别为0x9100和0x8100的VLAN标签，而在设备上配置的TPID字段值为0x9100，通过比较可以发现仅与运营商VLAN标签中的TPID字段值一致，所以设备将认为该帧仅承载了运营商VLAN标签，而把帧中的用户VLAN标签当成了数据部分。

另外，不同运营商的系统可能将QinQ帧外层VLAN标签的TPID设置为不同值。为实现与这些系统的兼容性，可以修改TPID值，使QinQ帧发送到公网时承载与特定运营商相同的 TPID 字段值，从而实现与该运营商设备之间的互操作性。但因以太网帧的TPID字段与不带VLAN标签的帧的“长度/类型/”（Length/Type/）字段位置相同，为避免在网络中转发和处理数据包时出现问题，不可将 TPID 值设置为表 7-8中各上层协议类型所对应的任意值。

表7-8 网络层协议类型及对应的十六进制值

协议类型	对应值
ARP	0x0806
RARP	0x8035
IP	0x0800
IPv6	0x86DD
PPPoE	0x8863/0x8864
MPLS	0x8847/0x8848
IPX/SPX	0x8137
LACP	0x8809
802.1x	0x888E
HGMP	0x88A7
设备保留	0xFFFD/0xFFFE/0xFFFF

7.3.4 QinQ映射

QinQ Mapping（QinQ映射）与本章后面将要介绍的VLAN映射类似，也是对数据帧中的VLAN标签进行替换，最根本的不同就是**QinQ**映射是在路由子接口上应用的，而**VLAN**映射是在物理端口上应用的 。

1．QinQ映射基本原理

QinQ映射发生在数据帧从入端口接收进来之后，从出端口转发出去之前。通过QinQ映射功能，子接口在向外发送本地 VLAN 的帧时，将帧中的本地 VLAN 标签替换成外部VLAN标签；在接收外部VLAN的帧时，又将帧中的外部VLAN标签替换成本地VLAN标签。可以看出，它的收、发过程中 VLAN 映射过程是相逆的。在实际组网中，QinQ映射功能可以将用户的 VLAN 标签映射为运营商的 VLAN 标签，从而起到屏蔽不同用户VLAN标签的作用。

QinQ映射功能一般部署在ISP网络边缘设备上，对用户侧上送的数据帧进行映射操作。将用户数据帧携带的VLAN标签映射为指定的VLAN标签后再接入公网。QinQ映射功能常应用于但不局限于以下场景。

（1）新局域网和老局域网部署的VLAN ID冲突，但是新局域网需要与老局域网互通。

（2）接入公网的各个局域网规划不一致，导致VLAN ID冲突。

（3）公网两端的VLAN ID规划不对称。

2．QinQ映射方式

目前，设备支持以下几种映射方式。

（1）1 to 1的映射方式。当部署QinQ映射功能设备上的子接口收到带有一层VLAN标签的数据帧时，将数据帧中携带的一层标签映射为用户指定的一层标签。发送帧的过程则相反。这其实与普通的VLAN映射的功能一样，只不过QinQ映射作用在路由子接口上。

（2）2 to 1的映射方式。当部署QinQ映射功能设备上的子接口收到带有两层标签的数据帧时，将数据

帧中携带的两层标签映射为用户指定的一层标签。发送帧的过程则相反。

下面以如图7-16所示的PC1向PC2发送帧为例进行介绍。具体流程如下。

（1）当PC1发送的数据帧到达启用了QinQ功能的Device1后被封装成双层标签（内层标签为10，外层标签为20）的VLAN帧。

（2）当配置了 2 to 1映射功能的Device2的GE1/0/1.1子接口接收到由Device1发来的 QinQ 帧后会把帧中原来的双层标签映射成一层标签 50（这个映射配置需要事先在Device2的GE1/0/1.1子接口上配置好），然后通过Device2的GE1/0/2端口向ISP网络发送数据帧并透传。

（3）当Device3上的GE1/0/2端口收到Device2发来的单层VLAN帧后，在转发到GE1/0/1.1子接口时再通过 2 to 1的QinQ映射功能将帧中的单层标签映射为的双层标签（内层标签为30、外层标签为40）的VLAN帧（这个映射配置也需要事先在Device3的GE1/0/1.1子接口上配置好），然后通过该子接口向Device4发送。

（4）当数据到了启用了 QinQ 功能的 Device4 后会去掉帧中的外层标签，形成单层标签（标签VLAN ID为 30）的VLAN帧，然后向PC2发送。

与PC2向PC1发送帧的流程同理。最终可实现PC1和PC2的互通，尽管它们原来所属的VLAN并不相同，但通过同一个外层VLAN也可实现互通。

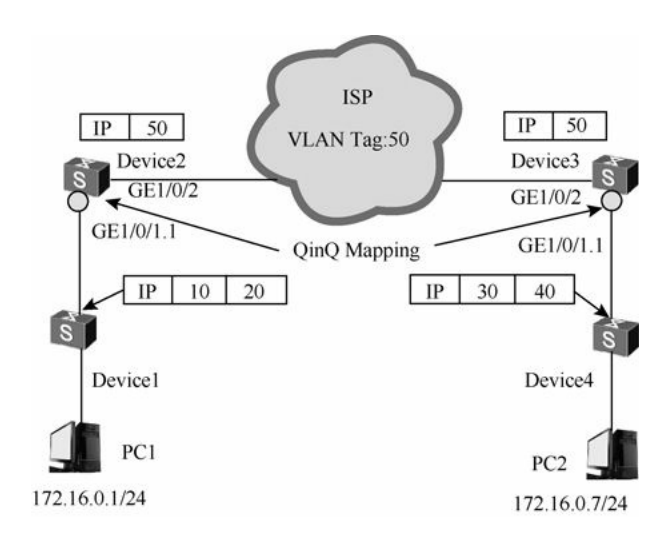


图7-16 QinQ映射应用示例

3．QinQ映射与VLAN映射的比较

在本章后面将介绍VLAN映射，在VLAN映射中也包括了以上QinQ的两种映射方式，它们的对比如表7-9所示。

表7-9 QinQ映射与VLAN映射的比较

映射类型	相同点	不同点
1 to 1	端口收到带有标签的数据帧后，将帧中的一层标签映射为用户指定的一层标签	QinQ 映射动作发生在子接口上，主要用于接入 VPLS（Virtual Private LAN Service，虚拟私有局域网服务）网络；VLAN 映射动作发生在干道物理端口上，主要用于通过 VLAN 转发的二层网络
2 to 1	入接口收到的帧带有两层标签	QinQ 映射动作发生在子接口上，子接口收到带有两层标签的数据帧后，将帧中的两层标签映射为用户指定的一层标签，主要用于接入 VPLS 网络；VLAN 映射动作发生在干道物理端口上，当干道端口收到带有两层标签的数据帧后，将帧中的外层标签映射为用户指定的一层标签，内层标签作为用户数据透传，主要用于通过 VLAN 转发的二层网络

7.4 基本QinQ配置与管理

为了使私网与公网有效分离，并最大限度地节省VLAN资源，可在设备端口上部署基本QinQ功能，新增外层802.1Q标签。其中内层标签用于标识内部网络（如企业网），外层标签用于标识外部网络（如运营商网络），从而最多可以提供 4 094×4 094个VLAN，并满足不同私网用户之间相同VLAN可以透明传输。

7.4.1 配置基本QinQ功能

基本 QinQ 功能是基于端口实现的，对于从端口进来的所有数据帧都加上同一个公网VLAN标签，实现用户数据帧在公网内转发。其实可以看成是一种基于端口的VLAN划分方式，具体的配置步骤如表7-10所示。

表7-10 基本QinQ功能配置步骤

步骤	命令	说明
1	**system-view** 例如：< HUAWEI > **system-view**	进入系统视图
2	**vlan** *vlan-id* 例如：[HUAWEI] **vlan** 2	创建外层 VLAN，并进入 VLAN 视图。参数 *vlan-id* 的取值范围为 1～4 094 的整数
3	**quit** 例如：[HUAWEI -VLAN2] **quit**	返回系统视图
4	**interface** *interface-type interface-number* 例如：[HUAWEI] **interface gigabitethernet 1/0/1**	键入要启用基本 QinQ 功能的交换机端口
5	**port link-type dot1q-tunnel** 例如：[HUAWEI-GigabitEthernet1/0/1]**port link-type dot1q-tunnel**	配置以上端口类型为 dot1q-tunnel（即 QinQ 类型）。QinQ 类型的端口用来连接其他交换机设备，并且能够处理携带双层标签的 VLAN 帧 缺省情况下，端口的链路类型为 Hybrid，改变端口类型前，要恢复端口只加入 VLAN1 的缺省配置，可用 **undo port link-type** 命令恢复端口类型为缺省的 Hybrid 类型
6	**port default vlan** *vlan-id* 例如：[HUAWEI-GigabitEthernet1/0/1]**port default vlan** 5	配置外层 VLAN 标签的 VLAN ID（即接口的缺省 VLAN）。参数 *vlan-id* 的取值范围为 1～4 094 的整数 缺省情况下，所有端口的缺省VLAN ID为1，可用**undo port default vlan** 命令删除端口配置的缺省 VLAN，恢复为缺省的 VLAN 1

以上这些命令已在第6章6.2节介绍基于端口的VLAN划分中作了详细介绍，在此不再赘述。配置完后可通过 **display current-configuration interface** interface-type interface-number命令查看指定端口上的QinQ配置。

7.4.2 配置外层VLAN标签的TPID值

如果需要实现不同厂商的设备互通，则端口的QinQ外层VLAN标签的协议类型标识（TPID）应配置为和该端口相连的设备能够识别的协议类型，需要配置外层VLAN标签的 TPID 值。但本节是可选配置任务，仅在网络中存在其他品牌设备时需要配置。而且本节介绍的配置同样适用于后面将要介绍的灵活QinQ配置和QinQ映射配置。

配置外层VLAN标签的TPID的方法很简单，只需在对应的交换机端口视图下使用 qinq protocol protocol-id命令配置即可。参数 protocol-id用来指定QinQ外层协议号，为4位16进制整数形式，取值范围是0x0600～0xFFFF，缺省值是0x8100。

配置本命令后，在入方向是对数据帧起到识别的作用，在出方向是对数据帧中的TPID进行修改或添加。但使用本命令配置的协议类型不能与一些特定的协议类型编号相同，否则会导致接口不能正确区分相应类型的协议报文，具体参见表7-8。

【示例 】配置GE1/0/1端口的QinQ报文外层VLAN标签的TPID值为0x9100。

<HUAWEI>system-view

[HUAWEI] interface gigabitethernet 1/0/1

[HUAWEI-GigabitEthernet1/0/1] qinq protocol 9100

7.4.3 配置对**Untagged**数据帧添加双层**VLAN**标签

通常如果要给数据帧打上双层标签，需要通过两台设备完成。配置本节介绍的功能后可实现通过一台设备给数据帧打上两层标签，方便了用户配置。也可实现当二层端口收到 Untagged 数据帧后根据实际业务或用户添加双层标签，达到区分业务或用户的目的。对Untagged数据帧添加双层VLAN标签的具体配置步骤如表7-11所示（本项配置任务同样适用于其他QinQ功能配置）。

表7-11 对Untagged数据帧添加双层VLAN标签的配置步骤

步骤	命令	说明
1	**system-view** 例如：< HUAWEI > **system-view**	进入系统视图
2	**vlan** *vlan-id* 例如：[HUAWEI] **vlan** 2	创建外层 VLAN，并进入 VLAN 视图。参数 *vlan-id* 的取值范围为 1～4 094 的整数
3	**quit** 例如：[HUAWEI -VLAN2] **quit**	返回系统视图
4	**interface** *interface-type interface-number* 例如：[HUAWEI] **interface gigabitethernet 1/0/1**	键入要启用基本 QinQ 功能的交换机端口
5	**port link-type hybrid** 例如：[HUAWEI-GigabitEthernet1/0/1] **port link-type hybrid**	配置以上端口类型为 Hybrid 类型，缺省情况下，端口的链路类型为 Hybrid，所以也可用 **undo port link-type** 命令恢复端口类型为缺省的 Hybrid 类型
6	**port hybrid Untagged vlan** *vlan-id* 例如：[HUAWEI-GigabitEthernet1/0/1] **port hybrid Untagged vlan** 2	把以上端口以不带标签方式添加到前面创建的外层 VLAN 中，参数 *vlan-id* 的取值范围为 1～4 094 的整数 缺省情况下，Hybrid 端口以 Untagged 方式加入 VLAN1，可用 **undo port hybrid Untagged vlan** *vlan-id* 命令删除所加入的外层 VLAN
7	**port vlan-stacking Untagged stack-vlan** *vlan-id1* **stack-inner-vlan** *vlan-id2* 例如：[HUAWEI-GigabitEthernet1/0/1] **port vlan-stacking Untagged stack-vlan** 2 **stack-inner-vlan** 5	对 Untagged 数据帧添加双层 VLAN 标签。命令中的参数说明如下。 （1）*vlan-id1*：指定对 Untagged 数据帧所添加的外层 VLAN 标签，且必须与上一步 **port hybrid Untagged vlan** *vlan-id* 命令中的 *vlan-id* 参数值一致 （2）*vlan-id2*：指定对 Untagged 数据帧所添加的内层 VLAN 标签，取值范围均为 1～4 094 的整数

（续表）

步骤	命令	说明
7	**port vlan-stacking Untagged stack-vlan** *vlan-id1* **stack-inner-vlan** *vlan-id2* 例如：[HUAWEI-GigabitEthernet1/0/1] **port vlan-stacking Untagged stack-vlan** 2 **stack-inner-vlan** 5	【说明】当端口的 PVID 不是缺省值 VLAN1 时，需要恢复端口的 PVID 为缺省值后才可以配置本命令，**目前只有 S7700、S9300、S9300E 和 S9700 系列的 E 系列和 F 系列单板支持** 另外，对 Untagged 数据帧添加双层 VLAN Tag 属于基于端口划分 VLAN 方式，同样遵守不同方式划分 VLAN 的优先级顺序，即基于端口划分 VLAN 的优先级最低，所以要使本命令配置生效，必须确保本地交换机上没有采用其他 VLAN 划分方式 缺省情况下，没有配置对 Untagged 数据帧添加双层 VLAN 标签，可用 **undo port vlan-stacking Untagged** 命令恢复缺省情况

完成配置后可使用 **display current-configuration interface** interface-type interface-number命令查看指定端口的外层VLAN标签配置。

【示例 】配置GE1/0/1端口对接收的Untagged报文添加双层VLAN标签，其中内层标签为200，外层标

签为100。

```
<HUAWEI>system-view
[HUAWEI] interface gigabitethernet 1/0/1
[HUAWEI-GigabitEthernet1/0/1] port hybrid Untagged vlan 100
[HUAWEI-GigabitEthernet1/0/1] port vlan-stacking Untagged stack-vlan100 stack-inner-vlan200
```

7.4.4 基本**QinQ**配置示例

如图7-17所示，网络中有两个企业，企业1（Enterprise1）和企业2（Enterprise2）各有两个分支。这两个企业的各办公地的企业网都分别和运营商网络中的 SwitchA 和SwitchB相连，且公网中存在其他厂商设备，其外层VLAN标签的TPID值为0x9100。

现要实现企业1和企业2独立划分VLAN，两者互不影响；各企业两分支机构之间的流量通过公网透明传输，相同企业之间可以互通，不同企业之间互相隔离。此时可通过配置基于端口的基本QinQ来实现以上需求。利用公网提供的VLAN 100使企业 1的两分支机构互通，利用公网提供的VLAN 200使企业 2的两分支机构互通（这里都要同时用到**QinQ**的添加外层标签和剥离外层标签的双重功能 ），同时通过QinQ功能可以实现不同企业之间的互相隔离，因为不同企业的数据帧中所封装的外层标签不同，外层标签对应的VLAN所加入的QinQ端口也不同。同时，通过在连接其他厂商设备的端口上配置修改QinQ外层VLAN标签的TPID值，来实现与其他厂商设备的互通。

1．配置思路分析

根据以上分析，可采用如下的思路进行本示例的基本QinQ配置。

（1）在 SwitchA和 SwitchB上均创建VLAN 100和VLAN 200，配置连接业务的GE1/0/1和GE1/0/2端口为QinQ类型，并分别加入对应的外层VLAN，同时可实现在发送数据帧时去掉帧中的外层VLAN标签的功能。

（2）配置SwitchA和SwitchB上连接公网的GE1/0/3端口为Trunk类型，然后同时允许VLAN 100和VLAN 200的数据帧通过。

（3）在SwitchA和SwitchB连接公网的GE1/0/3端口上配置外层VLAN标签的TPID值，实现与其他厂商设备的互通。

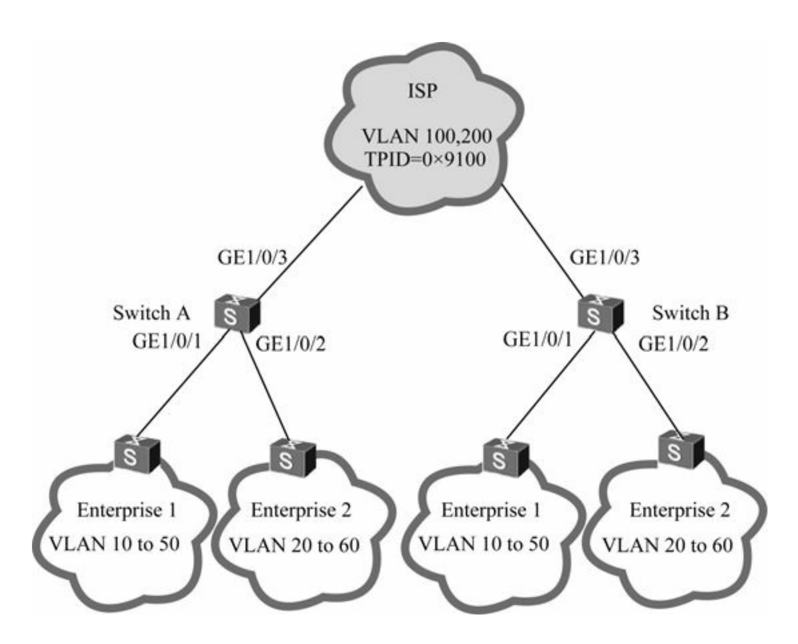


图7-17 基本QinQ配置示例

2．具体配置步骤

（1）在SwitchA和SwitchB上均创建公网VLAN 100和VLAN 200。

SwitchA上的配置：

```
<HUAWEI>system-view
[HUAWEI] sysname SwitchA
[SwitchA] vlan batch 100 200
```

SwitchB上的配置：

```
<HUAWEI>system-view
[HUAWEI] sysname SwitchB
[SwitchB] vlan batch 100 200
```

（2）把SwitchA和SwitchB的GE1/0/1、GE1/0/2端口上配置为QinQ类型，并加入对应的公网 VLAN 中。这样就在从端口上接收到的数据帧上添加对应的外层 VLAN 标签，同时允许对应的外层VLAN通过。

SwitchA上的配置：

```
[SwitchA] interface gigabitethernet 1/0/1
[SwitchA-GigabitEthernet1/0/1] port link-type dot1q-tunnel
[SwitchA-GigabitEthernet1/0/1] port default vlan 100
[SwitchA-GigabitEthernet1/0/1] quit
[SwitchA] interfacegigabitethernet 1/0/2
[SwitchA-GigabitEthernet1/0/2] port link-type dot1q-tunnel
[SwitchA-GigabitEthernet1/0/2] port default vlan 200
[SwitchA-GigabitEthernet1/0/2] quit
```

SwitchB上的配置：

```
[SwitchB] interface gigabitethernet 1/0/1
[SwitchB-GigabitEthernet1/0/1] port link-type dot1q-tunnel
[SwitchB-GigabitEthernet1/0/1] port default vlan 100
[SwitchB-GigabitEthernet1/0/1] quit
[SwitchB] interfacegigabitethernet 1/0/2
[SwitchB-GigabitEthernet1/0/2] port link-type dot1q-tunnel
[SwitchB-GigabitEthernet1/0/2] port default vlan 200
[SwitchB-GigabitEthernet1/0/2] quit
```

（3）配置SwitchA和SwitchB连接公网侧的GE1/0/3端口为Trunk类型，并允许VLAN 100和VLAN 200这两个公网VLAN通过，因为每台交换机都连接了企业 1和企业2，也就有对应的两种封装了不同外层标签的数据帧进入。

SwitchA上的配置：

```
[SwitchA] interfacegigabitethernet 1/0/3
[SwitchA-GigabitEthernet1/0/3] port link-type trunk
[SwitchA-GigabitEthernet1/0/3] port trunk allow-pass vlan 100 200
[SwitchA-GigabitEthernet1/0/3] quit
```

SwitchB上的配置：

[SwitchB] interfacegigabitethernet 1/0/3

[SwitchB-GigabitEthernet1/0/3] port link-type trunk

[SwitchB-GigabitEthernet1/0/3] port trunk allow-pass vlan 100 200

[SwitchB-GigabitEthernet1/0/3] quit

（4）在SwitchA和SwitchB的GE1/0/3端口上均配置外层VLAN标签的TPID值为9100，以便与其他品牌设备兼容。

SwitchA上的配置：

[SwitchA] interfacegigabitethernet 1/0/3

[SwitchA-GigabitEthernet1/0/3] qinq protocol 9100

SwitchB上的配置：

[SwitchB] interface gigabitethernet 1/0/3

[SwitchB-GigabitEthernet1/0/3] qinq protocol 9100

完成以上配置后，因为QinQ帧在对方启用了QinQ功能的端口后又会剥离相应的外层标签，所以从企业1的一处分支机构内的任意VLAN的一台PC可以ping得通企业1的另一处分支机构的相同VLAN内的PC，从企业2的一处分支机构内的任意VLAN的一台PC可以ping得通企业2的另一处分支机构的相同VLAN内的PC，实现相同企业的相同VLAN间的用户之间互通。而企业1中的任意VLAN中的PC都不能ping得通企业2任意VLAN内的PC，实现企业1和企业2用户之间的相互隔离。

7.5 灵活QinQ配置与管理

灵活QinQ是对QinQ的一种更灵活的实现，它是基于端口与VLAN相结合的方式实现的，可以对进入同一端口的数据帧依据帧中原来的内层VLAN ID的不同来添加不同的外层VLAN标签，当然它与前面介绍的基本QinQ一样，在发送数据帧时也会剥离帧中的外层**VLAN**标签 。这样就可以实现对不同用户和不同业务的区分。

下面根据7.3.2节介绍的三种不同的灵活QinQ封装类型介绍不同的灵活QinQ配置方法。

7.5.1 配置基于VLAN ID的灵活QinQ

基于VLAN ID的灵活QinQ功能可实现端口在接收到数据帧后，依据帧中不同内层VLAN ID添加不同的外层VLAN标签。它与7.4节介绍的基于端口的基本QinQ不一样，因为不能像7.4.1节介绍的那样直接把对应的交换机端口配置为QinQ类型端口，而是通 过一条**qinq vlan-translation enable** 命令在端口上使能VLAN转换功能。具体配置步骤如表7-12所示。

注意

配置基于VLAN ID的灵活QinQ功能的端口类型必须是**Hybrid**类型， 且只在入 方向上生效，同时必须以Untagged方式加入添加后的外层VLAN中（当然此外层VLAN必须事先创建），这样就可使得该端口在发送数据帧时去掉帧中的外层 VLAN，实现外层标签的剥离目的。在端口学习MAC地址时，所学习到的MAC地址是QinQ数据帧外层VLAN的MAC地址，与数据帧原来的内层VLAN无关。

表7-12 基于VLAN ID的灵活QinQ的配置步骤

步骤	命令	说明
1	**system-view** 例如：< HUAWEI > **system-view**	进入系统视图
2	**vlan** *vlan-id* 例如：[HUAWEI] **vlan** 2	创建外层 VLAN，并进入 VLAN 视图。参数 *vlan-id* 的取值范围为 1～4 094 的整数
3	**quit** 例如：[HUAWEI -VLAN2] **quit**	返回系统视图
4	**Interface** *interface-type interface-number* 例如：[HUAWEI] **interface gigabitethernet 1/0/1**	键入要启用基本 QinQ 功能的交换机端口
5	**port link-type hybrid** 例如：[HUAWEI-GigabitEthernet1/0/1]**port link-type hybrid**	配置以上端口类型为 Hybrid 类型，其他说明参见表 7-11 第 5 步
6	**port hybrid Untagged vlan** *vlan-id* 例如：[HUAWEI-GigabitEthernet1/0/1]**port hybrid Untagged vlan** 2	把以上 Hybrid 端口以 Untagged 方式加入到外层 VLAN 中，参数 *vlan-id* 用来指定前面创建的外层 VLAN，取值范围为 1～4 094 的整数
7	**qinq vlan-translation enable** 例如：[HUAWEI-GigabitEthernet1/0/1]**qinq vlan-translation enable**	使能以上端口的 VLAN 转换功能。仅在端口使能了 VLAN 转换功能后才可以在端口上配置 VLAN 映射和灵活 QinQ 功能。**但本命令在 S7700、S9300 和 S9700 系列中不支持，也就不需要配置此命令** 缺省情况下，没有使能接口 VLAN 转换功能，可用 **undo qinq vlan-translation enable** 命令取消端口的 VLAN 转换功能
8	**port vlan-stacking vlan** *vlan-id1*[**to** *vlan-id2*] **stack-vlan** *vlan-id3* [**remark-8021p** *8021p-value*] 例如：[HUAWEI-GigabitEthernet1/0/1] **port vlan-stacking vlan** 1 **to** 4 **stack-vlan** 3	配置灵活 QinQ，也即 VLAN Stacking 功能。命令中的参数说明如下。 （1）*vlan-id1* [**to** *vlan-id2*]：指定要添加由参数 *vlan-id3* 指定的外层标签的内层 VLAN ID 范围，其中 *vlan-id1* 表示起始 VLAN ID；**to** *vlan-id2* 表示结束 VLAN ID，取值范围均为 1～4 094 的整数 **此时要特别注意，添加不同外层 VLAN 的内层 VLAN ID 范围绝对不能重叠，或者交叉，否则就会使端口无法正确添加外层 VLAN 标签** （2）*vlan-id3*：指定添加的外层标签对应的 VLAN ID，取值范围为 1～4 094 的整数

（续表）

步骤	命令	说明
8	**port vlan-stacking vlan** *vlan-id1*[**to** *vlan-id2*] **stack-vlan** *vlan-id3* [**remark-8021p** *8021p-value*] 例如：[HUAWEI-GigabitEthernet1/0/1] **port vlan-stacking vlan** 1 **to** 4 **stack-vlan** 3	（3）*8021p-value*：可选参数，重新标记添加外层标签后帧的 802.1p 优先级，取值范围为 0～7，值越大优先级越高。缺省情况下，对于 S7700、S9300、S9300E 和 S9700 系列中的 SA 单板的外层 VLAN 优先级为 0，其他单板的外层 VLAN 优先级与内层 VLAN 优先级保持一致；对于其他情况下，外层 VLAN 的 802.1p 优先级与内层 VLAN 的 802.1p 优先级保持一致 缺省情况下，没有配置 VLAN Stacking 功能，可用 **undo port vlan-stacking vlan** *vlan-id1* [**to** *vlan-id2*] [**stack-vlan** *vlan-id3*] 命令取消对应的 VLAN Stacking 功能

配置完成后，使用**display current-configuration interface** interface-type interface-number命令查看端口的灵活QinQ配置。

【示例 】配置GE0/0/1端口的灵活QinQ功能，对用户VLAN标签为VLAN 10～VLAN 13的数据帧添加外层VLAN标签100。

<HUAWEI>system-view

[HUAWEI] interface gigabitethernet 0/0/1

[HUAWEI-GigabitEthernet0/0/1] qinq vlan-translation enable

[HUAWEI-GigabitEthernet0/0/1] port hybrid Untagged vlan 100

[HUAWEI-GigabitEthernet0/0/1] port vlan-stacking vlan 10 to 13 stack-vlan 100

7.5.2 基于VLAN ID的灵活QinQ配置示例

本示例拓扑结构如图7-18所示，PC上网用户（假设用户VLAN ID范围为100～200）和VoIP用户（假设用户VLAN ID范围为 300～400）通过SwitchA和SwitchB接入，并分别以VLAN 2和VLAN 3通过运营商网络（Carrier Network）。现要求PC上网用户和VoIP用户通过运营商网络实现互相通信。

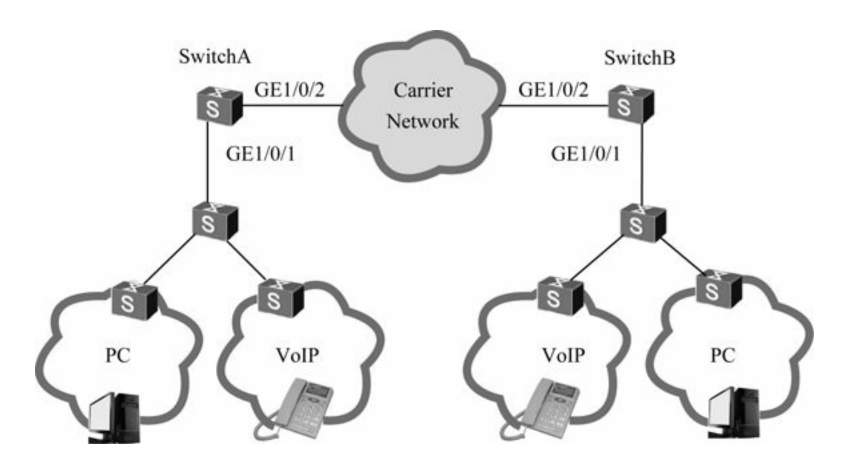


图7-18 基于VLAN ID的灵活QinQ配置示例的拓扑结构

1．配置思路分析

本示例的基本配置思路很简单，具体如下。

（1）在SwitchA和SwitchB上创建所需的外层VLAN 2和VLAN 3（用户的内层VLAN可不创建），并将GE1/0/1端口配置为Untagged方式的Hybrid类型端口，然后都加入外层VLAN 2和VLAN 3中。

（2）在SwitchA和SwitchB的GE1/0/1端口上配置灵活QinQ功能，以实现在接收数据帧时依据其内层VLAN标签添加对应的外层VLAN标签，在发送QinQ帧时去掉对应的外层VLAN标签，以实现正常的流量转发。

2．具体配置步骤

（1）在SwitchA和SwitchB上创建外层VLAN 2和VLAN 3，因为这两台交换机都同时连接了PC上网用户和VoIP用户。

SwitchA上的配置：

```
<HUAWEI>system-view
[HUAWEI] sysname SwitchA
[SwitchA] vlan batch 2 3
```

SwitchB上的配置：

```
<HUAWEI>system-view
[HUAWEI] sysname SwitchB
[SwitchB] vlan batch 2 3
```

（2）在SwitchA和SwitchB上行连接运营商网络的Untagged方式Hybrid类型GE1/0/1端口上配置基于VLAN ID的灵活QinQ功能，添加双层VLAN标签。但要注意，不同用户的内层**VLAN ID**绝对不能重叠和交叉。

SwitchA上的配置：

```
[SwitchA] interface gigabitethernet 1/0/1
```

[SwitchA-GigabitEthernet1/0/1] port link-type hybrid

[SwitchA-GigabitEthernet1/0/1] port hybrid Untagged vlan 2 3 !---指定GE1/0/1端口以不带标签方式加入VLAN 2和VLAN 3

[SwitchA-GigabitEthernet1/0/1] qinq vlan-translation enable !---如果交换机是 S7700、S9300或 S9700系列，则不需要配置此命令，下同

[SwitchA-GigabitEthernet1/0/1] port vlan-stacking vlan 100 to 200 stack-vlan 2 !---为PC上网用户添加外层VLAN标签为2

[SwitchA-GigabitEthernet1/0/1] port vlan-stacking vlan 300 to 400 stack-vlan 3 !---为VoIP网用户添加外层VLAN标签为3

[SwitchA-GigabitEthernet1/0/1] quit

SwitchB上的配置：

[SwitchB] interface gigabitethernet 1/0/1

[SwitchB-GigabitEthernet1/0/1] port link-type hybrid

[SwitchB-GigabitEthernet1/0/1] port hybrid Untagged vlan 2 3

[SwitchB-GigabitEthernet1/0/1] qinq vlan-translation enable

[SwitchB-GigabitEthernet1/0/1] port vlan-stacking vlan 100 to 200 stack-vlan 2

[SwitchB-GigabitEthernet1/0/1] port vlan-stacking vlan 300 to 400 stack-vlan 3

[SwitchB-GigabitEthernet1/0/1] quit

（3）配置SwitchA和SwitchB连接运营商网络的上行GE1/0/2端口为Trunk类型，并同时加入外层VLAN 2和VLAN 3中。

SwitchA上的配置：

[SwitchA] interfacegigabitethernet1/0/2

[SwitchA-GigabitEthernet1/0/2] port link-type trunk

[SwitchA-GigabitEthernet1/0/2] port trunk allow-pass vlan 2 3

[SwitchA-GigabitEthernet1/0/2] quit

SwitchB上的配置：

[SwitchB] interface gigabitethernet 1/0/2

[SwitchB-GigabitEthernet1/0/2] port link-type trunk

[SwitchB-GigabitEthernet1/0/2] port trunk allow-pass vlan 2 3

[SwitchB-GigabitEthernet1/0/2] quit

最后来验证以上配置结果，可通过display current-configuration interface命令查看SwitchA上GE1/0/1和GE1/0/2端口上的配置信息，看是否正确。具体如下。

<SwitchA>display current-configuration interfacegigabitethernet1/0/1

#

interface GigabitEthernet1/0/1

port hybrid Untagged vlan 2 to 3

qinq vlan-translation enable

port vlan-stacking vlan 100 to 200 stack-vlan 2

port vlan-stacking vlan 300 to 400 stack-vlan 3

#

```
return
<SwitchA>display current-configuration interfacegigabitethernet1/0/2
#
interface GigabitEthernet1/0/2
port link-type trunk
port trunk allow-pass vlan 2 to 3
#
return
```

可以用同样的方法查看SwitchB上的GE1/0/1和GE1/0/2端口上的配置信息，以验证配置是否正确。如果SwitchA、SwitchB上配置正确，则可实现PC上网用户通过运营商网络互相通信，VoIP用户也可以通过运营商网络互相通信。

7.5.3 配置基于802.1p优先级的灵活QinQ

基于802.1p优先级（也就是通常所说的VLAN优先级）的灵活QinQ功能可以根据进入端口的数据帧的802.1p优先级和VLAN ID灵活地添加外层VLAN标签，优先保证重要用户的正常通信。仅在**S7700**、**S9300**和**S9700**系列的中**E**子系列和**F**子系列单板支持 。具体的配置步骤如表7-13所示。

注意

基于802.1p优先级的灵活QinQ功能仅对入方向的数据帧生效，且配置此功能的端口的类型必须为Trunk或Hybrid类型。

表7-13 基于802.1p优先级的灵活QinQ的配置步骤

步骤	命令	说明
1	**system-view** 例如：<HUAWEI> **system-view**	进入系统视图
2	**Interface** *interface-type interface-number* 例如：[HUAWEI]**interface gigabitethernet 0/0/1**	键入要配置1 to 1 VLAN映射的交换机端口，进入接口视图
3	**port hybrid Untagged vlan** *vlan-id* 例如：[HUAWEI-GigabitEthernet0/0/1]**port hybrid Untagged vlan** 20	配置以上入端口以不带标签方式加入外层VLAN(此VLAN必须事先创建好)。参数 *vlan-id* 指定要加入的外层VLAN的ID，取值范围为1～4 094的整数 缺省情况下，Hybrid端口以Untagged方式加入VLAN1，可用 **undo port hybrid vlan** *vlan-id* 命令删除Hybrid类型端口加入外层VLAN

（续表）

步骤	命令	说明
4	**port vlan-stacking 8021p** *8021p-value* **stack-vlan** *vlan-id* 例如：[HUAWEI-GigabitEthernet0/0/1]**port vlan-stacking 8021p** 5 **stack-vlan** 20	（二选一）配置以上端口基于 802.1p 优先级的灵活 QinQ 功能。命令中的参数说明如下。 （1）*8021p-value*：指定添加外层 VLAN 标签后帧的优先级，取值范围是 1～7 的整数，值越大优先级越高 （2）*vlan-id*：指定添加的外层 VLAN 的 ID，取值范围为 1～4 094 的整数 缺省情况下，没有配置 VLAN Stacking 功能，可用 **undo port vlan-stacking 8021p** *8021p-value1* [**stack-vlan** *vlan-id*] 命令删除端口基于指定 802.1p 优先级的 VLAN Stacking 功能
5	**port vlan-stacking vlan** *vlan-id1* [**to** *vlan-id2*] **8021p** *8021p-value1*[**to** *8021p-value2*] **stack-vlan** *vlan-id3*[**remark-8021p** *8021p-value3*] 例如：[HUAWEI-GigabitEthernet0/0/1]**port vlan-stacking vlan** 100 **8021p** 5 **stack-vlan** 20 **remark-8021p** 1	（二选一）配置接口基于 VLAN 和 802.1p 优先级的灵活 QinQ 功能。命令中的参数说明如下。 （1）*vlan-id1* [**to** *vlan-id2*]：指定要添加外层 VLAN 标签的帧中 VLAN ID 范围，其中 *vlan-id1* 表示起始 VLAN ID，可选参数 **to** *vlan-id2* 表示结束 VLAN ID，取值范围均为 1～4 094 的整数，但 *vlan-id2* 的取值必须大于 *vlan-id1* 的取值，它和 *vlan-id1* 共同确定一个范围 （2）*8021p-value1* [**to** *8021p-value2*]：指定添加外层 VLAN 标签的帧中 802.1p 优先级的取值范围，其中 *8021p-value1* 表示 802.1p 优先级取值范围的下限，**to** *8021p-value2* 表示 802.1p 优先级取值范围的上限 （3）*vlan-id3*：指定要添加的外层标签 VLAN ID，取值范围为 1～4 094 的整数 （4）*8021p-value3*：可选参数，重标记添加外层标签后的 VLAN 帧的 802.1p 优先级。通过本可选参数的设置，可实现端口在接收到带 VLAN 标签的数据帧后，将帧中的 802.1p 优先级修改为用户配置的 802.1p 优先级值 缺省情况下，交换机端口下没有配置对数据帧中携带的 VLAN 标签进行外层 VLAN 标签添加操作，可用 **undo port vlan-stacking vlan** *vlan-id1* [**to** *vlan-id2*] [**8021p** *8021p-value1* [**to** *8021p-value2*]] [**map-vlan** *vlan-id3*]命令取消对应交换机端口基于 VLAN+802.1p 优先级的 VLAN Stacking 功能

注意

如果端口上同时配置了 **port vlan-stacking vlan** vlan-id1 [**to** vlan-id2] **8021p** 8021p-value1 [to 8021p-value2] **stack-vlan** vlan-id3命令（没有指定 **remark- 8021p** 8021p-value3参数）与 **port vlan-mapping vlan 8021p** 8021p-value1 [**to** 8021p-value2] **map-vlan** vlan-id3命令（指定了映射后 802.1p的优先级）。此时，启用灵活QinQ功能后帧的 802.1p优先级按 **port vlan-mapping vlan 8021p** 命令配置生效。

如果在入端口上创建了DiffServ（差分服务）域，并配置VLAN帧的802.1p优先级映射，则此时交换机的内部优先级（是指不同的服务级别）配置就会与帧中原来的802.1p优先级不一样。这时就需要在出端口上再次配置802.1p优先级映射，重新恢复帧中原来 的优先级。具体配置步骤如表7-14所示。

表7-14 在出端口上配置802.1p优先级映射的配置步骤

步骤	命令	说明
1	**system-view** 例如：< HUAWEI > **system-view**	进入系统视图
2	**diffserv domain** *ds-domain-name* 例如：[HUAWEI] **diffserv domain** d1	创建 DiffServ 域并进入 DiffServ 域视图。参数 *ds-domain-name* 用来指定 DiffServ 域的名称，为长度为 1～31 的字符串，但不支持空格，不区分大小写，不能为"n"、"no"、"non"、"none" 缺省情况下，系统预定义了一个名为 default 的 DiffServ 域，可用 **undo diffserv domain** *ds-domain-name* 命令删除指定的 DiffServ 域
3	**8021p-outbound** *service-class* { **green** | **yellow** | **red** } **map** *8021p-value* 例如：[HUAWEI-dsdomain-d1]**8021p-outbound** af1 **yellow map** 2	将 DiffServ 域中端口出方向上 VLAN 数据帧的内部优先级映射为指定的 802.1p 优先级。命令中的参数和选项说明如下。 （1）*service-class*：指定 PHB 行为，取值可以为 BE、AF1～AF4、EF、CS6 或 CS7（不区分大小写） （2）**green**：多选一选项，指定数据帧标记的颜色为绿色 （3）**yellow**：多选一选项，指定数据帧标记的颜色为黄色 （4）**red**：多选一选项，指定数据帧标记的颜色为红色 （5）*8021p-value*：指定 VLAN 数据帧中原来的 802.1p 优先级值，取值范围是 0～7 的整数，值越大优先级越高 【说明】当对 VLAN 数据帧进行了 QoS 调度之后，可以通过本命令配置 DiffServ 域中数据帧的 PHB 行为/颜色到 802.1p 优先级之间的映射。将 DiffServ 域绑定到数据帧的出接口后，下游设备将根据数据帧的 802.1p 优先级进行调度。这方面请参见本书第 10 章 可用 **undo 8021p-outbound** [*service-class* { **green** | **yellow** | **red** }]命令恢复缺省的映射关系。如果没有指定参数 *service-class* 和颜色选项，将恢复所有服务等级和数据帧颜色与对应的 802.1p 值的缺省配置
4	**quit** 例如：[HUAWEI-dsdomain-d1] **quit**	返回系统视图
5	**interface** *interface-type interface-number* 例如：[HUAWEI] **interface** gigabitethernet 1/0/1	键入出端口，进入接口视图
6	**port link-type** {**hybrid** | **trunk** } 例如：[HUAWEI-GigabitEthernet1/0/1] **port link-type hybrid**	配置以上出端口为 Hybrid 或者 Trunk 类型（基于 802.1p 优先级的灵活 QinQ 功能可以是 Hybrid 或者 Trunk 类型）。缺省情况下，端口类型为 Hybrid，可用 **undo port link-type** 命令恢复端口为缺省的 Hybrid 类型
7	**port hybrid tagged vlan** *vlan-id* 例如：[HUAWEI-GigabitEthernet1/0/1]**port hybrid tagged vlan** 20 或 **port trunk allow-pass vlan** *vlan-id* 例如：[HUAWEI-GigabitEthernet1/0/1] **port trunk allow-pass vlan** 20	配置以上 Hybrid 出端口以带标签方式加入指定的外层 VLAN，或者配置以上 Trunk 类型出端口允许指定的外层 VLAN 通过，参数的取值范围为 1～4 094 的整数，要与表 7-15 中所添加的外层标签 VLAN ID 一致

（续表）

步骤	命令	说明
8	**trust upstream** *ds-domain-name* 例如：[HUAWEI-GigabitEthernet1/0/1] **trust upstream** d1	在以上 Hybrid 或 Trunk 类型端口上应用前面的 DiffServ 域 VLAN 优先级映射配置。缺省情况下，端口上不应用任何 DiffServ 域，可用 **undo trust upstream** 命令恢复缺省配置 **本命令为覆盖式命令，即在同一端口视图下多次执行该命令配置后，按最后一次配置生效**。但如果要修改端口下应用的 DiffServ 域，必须先执行 **undo trust upstream** 命令删除已应用的 DiffServ 域，再执行 **trust upstream** 命令重新应用新的 DiffServ 域

【示例 1】在GE1/0/1端口上配置基于802.1p优先级的灵活QinQ（即VLAN Stacking）功能，对VLAN ID为100、802.1p优先级为5的帧添加外层标签200，并重标记帧的802.1p优先级为1。

<HUAWEI>system-view

[HUAWEI] interface gigabitethernet 1/0/1

[HUAWEI-GigabitEthernet1/0/1] port vlan-stacking vlan 100 8021p 5 stack-vlan 200 remark-8021p 1

【示例 2】在DiffServ域ds1中配置端口出方向上PHB行为是AF1的黄色VLAN报文对应802.1p优先级为

2。

<HUAWEI>system-view

[HUAWEI] diffserv domain ds1

[HUAWEI-dsdomain-ds1] 8021p-outbound af1 yellow map 2

【示例 3】在GE1/0/1端口上应用DiffServ域ds1。

<HUAWEI>system-view

[HUAWEI] interface gigabitethernet 1/0/1

[HUAWEI-GigabitEthernet1/0/1] trust upstreamds1

完成以上配置后，可在入或出端口视图下执行 display this命令查看该端口上基于802.1p优先级的灵活QinQ的配置信息。

7.5.4 配置基于流策略的灵活QinQ

这里所说的“流策略”是指将流分类和流行为关联后形成的完整的QoS策略（有关QoS策略方面的基础知识具体参见本书第 10章）。用户可以根据数据帧中的VLAN ID进行流分类，然后将流分类与某种流行为关联，对符合流分类的数据帧进行相应的处理（添加外层VLAN标签），从而实现灵活QinQ功能。基于流策略的灵活QinQ功能能够针对业务类型提供差别服务。

根据以上分析可以得出，基于流策略的灵活QinQ配置包括以下4个基本任务。

（1）定义流分类：针对数据帧中的内层VLAN ID进行分类。

（2）定义流行为：根据不同的内层VLAN ID添加不同的外层VLAN ID。

（3）创建QoS策略，将以上定义的流分类与流行为进行关联。

（4）在端口（必须是Hybrid类型端口）的入方向上应用以上创建QoS策略。

基于流策略的灵活QinQ的具体配置步骤如表7-15所示。但应用流策略的端口只能是不带标签的**Hybrid**类型 。

表7-15 基于流策略的灵活QinQ的配置步骤

配置任务	步骤	命令	说明
定义流分类	1	**system-view** 例如：< HUAWEI > **system-view**	进入系统视图
	2	**traffic classifier** *classifier-name* 例如：[HUAWEI] **traffic classifier** c1	创建流分类并进入流分类视图。参数 *classifier-name* 用来指定流分类名称，为 1～31 个字符，**且需以字母开头，不支持空格，区分大小写** 缺省情况下，系统没有定义任何流分类，可用 **undo traffic classifier** *classifier-name* 命令删除指定的流分类
	3	**if-match vlan-id** *start-vlan-id* [**to** *end-vlan-id*] 例如：[HUAWEI-classifier-c1]**if-match vlan-id** 2	用来在流分类中创建基于 VLAN ID 进行分类的匹配规则，指定数据帧的内层 VLAN ID。其他说明参见表 7-15 中的第 3 步
	4	**quit** 例如：[HUAWEI-classifier-c1]**quit**	退出流分类视图，返回系统视图
定义流行为	5	**traffic behavior** *behavior-name* 例如：[HUAWEI] **traffic behavior** b1	创建流行为并进入流行为视图。参数 *behavior-name* 用来指定流行为名称，为 1～31 个字符，**且需以字母开头，不支持空格，区分大小写** 缺省情况下，系统未创建任何流行为，可用 **undo traffic behavior** *behavior-name* 命令删除指定的流行为
	6	**nest top-most vlan-id** *vlan-id* 例如：[HUAWEI-behavior-b1]**nest top-most vlan-id** 200	在流行为中配置创建外层 VLAN 标签的动作。参数 *vlan-id* 用来指定创建的外层标签的 VLAN ID 值（**必须已在本地交换机上创建**），取值范围为 1～4 094 的整数 缺省情况下，流行为中没有配置创建外层 VLAN 标签的动作，可用 **undo nest** 命令在流行为中取消创建外层 VLAN 标签的动作
	7	**quit** 例如：[HUAWEI-behavior-b1] **quit**	退出流行为视图，返回系统视图
创建 QoS 策略，关联流分类与流行为	8	**traffic policy** *policy-name* [HUAWEI] **traffic policy** p1	创建流策略并进入流策略视图。参数 *policy-name* 用来指定创建的流策略名称，为 1～31 个字符，**且需以字母开头，不支持空格，区分大小写** 缺省情况下，系统没有创建任何流策略，可用 **undo traffic policy** *policy-name* 命令删除指定的流策略
	9	**Classifier** *classifier-name* **behavior** *behavior-name* 例如：[HUAWEI-trafficpolicy-p1] **classifier** c1 **behavior** b1	将以上定义的流分类与指定的流行为进行绑定，组成流策略。创建流策略后，必须在流策略视图下将流分类和相应的流行为关联起来，即绑定流分类和流行为，使流策略具有实际内容，该策略的应用才有意义 缺省情况下，流策略中没有绑定流分类和流行为，可用 **undo classifier** *classifier-name* 命令在流策略中取消流分类和流行为的绑定
	10	**quit** 例如：[HUAWEI-dsdomain-ds1] **quit**	退出流策略视图，返回系统视图

（续表）

配置任务	步骤	命令	说明
在交换机端口入方向上应用 QoS 策略	11	**Interface** *interface-type interface-number* 例如：[HUAWEI] **interface gigabitethernet 1/0/1**	键入要应用流策略的交换机端口，进入接口视图
	12	**port link-type hybrid** 例如：[HUAWEI-GigabitEthernet1/0/1] **port link-type hybrid**	配置以上端口的类型为 Hybrid
	13	**port hybrid Untagged vlan** { { *vlan-id1* [**to** *vlan-id2*] }&<1-10> | **all** } 例如：[HUAWEI-GigabitEthernet1/0/1] **port hybrid Untagged vlan** 2 3	把以上 Hybrid 端口以不带标签方式加入指定的外层 VLAN 中，具体命令参数本书第 6 章已有详细介绍，不再赘述
	14	**traffic-policy** *policy-name* **inbound** 例如：[HUAWEI-GigabitEthernet1/0/1] **traffic-policy p1 inbound**	在以上 Hybrid 端口的入方向应用流策略 缺省情况下，交换机端口上没有应用任何流策略，可用 **undo traffic-policy** [*policy-name*] **inbound** 命令取消在端口上应用流策略

完成配置后，可使用**display current-configuration** 命令查看基于流策略的灵活QinQ配置。

【示例 1】为流行为b1配置创建外层VLAN标签100的动作。

<HUAWEI>system-view

[HUAWEI] traffic behaviorb1

[HUAWEI-behavior-tb] nest top-most vlan-id 100

【示例 2】在新创建的流策略p1中配置流分类c1关联流行为b1，然后在GE1/0/1端口入方向上应用该流策略。

<HUAWEI>system-view

[HUAWEI] traffic policy p1

[HUAWEI-trafficpolicy-p1] classifier c1 behaviorb1

[HUAWEI-trafficpolicy-p1] quit

[HUAWEI] interface gigabitethernet 1/0/1

[HUAWEI-GigabitEthernet1/0/1] traffic-policy p1 inbound

7.5.5 基于流策略的灵活QinQ配置示例

7.5.2 节图 7-18 所示的示例也可采用基于策略的灵活 QinQ 配置方法，实现连接在SwitchA和SwitchB上的PC上网用户（内层标签为100～200）和VoIP用户（内层标签为 300～400）分别以VLAN 2和VLAN 3通过运营商互相通信。

1．配置思路分析

本示例采用基于流策略来配置灵活 QinQ 功能时关键是要创建正确的流分类和流行为，【这里的流分类就是基于帧中的内层VLAN ID进行的分类，流行为就是对不同范围的内层VLAN ID添加不同的外层VLAN。最后只需创建一个QoS策略，把以上流分类和流行为关联起来，并应用到对应的交换机端口上。具体配置思路如下。

（1）在SwitchA和SwitchB上创建所需的外层VLAN，然后定义基于内层VLAN ID的流分类，定义对应的流行为。

（2）在 SwitchA 和 SwitchB 下行端口为不带标签的 Hybrid 类型并加入所需的外层VLAN中，然后应用流策略来实现灵活QinQ功能。

（3）在 SwitchA 和 SwitchB 连接运营商网络的端口上配置为 Trunk 或者带标签的Hybrid类型，并允许

所有的外层VLAN通过。

2．具体配置步骤

（1）在SwitchA和SwitchB上创建所需的外层VLAN 2和VLAN 3。

SwitchA上的配置：

<HUAWEI>system-view

[HUAWEI] sysname SwitchA

[SwitchA] vlan batch 2 3

SwitchB上的配置：

<HUAWEI>system-view

[HUAWEI] sysname SwitchB

[SwitchB] vlan batch 2 3

（2）在SwitchA和SwitchB上配置流策略。

SwitchA上的配置：

[SwitchA] traffic classifier c1

[SwitchA-classifier-c1] if-match vlan-id 100 to 200

[SwitchA-classifier-c1] quit

[SwitchA] traffic behavior b1

[SwitchA-behavior-b1] nest top-most vlan-id 2

[SwitchA-behavior-b1] quit

[SwitchA] traffic classifier c2

[SwitchA-classifier-c2] if-match vlan-id 300 to 400

[SwitchA-classifier-c2] quit

[SwitchA] traffic behaviorb2

[SwitchA-behavior-b2] nest top-most vlan-id 3

[SwitchA-behavior-b2] quit

[SwitchA] traffic policy p1

[SwitchA-trafficpolicy-p1] classifier c1 behaviorb1

[SwitchA-trafficpolicy-p1] classifier c2 behaviorb2

[SwitchA-trafficpolicyp1] quit

SwitchB上的配置：

[SwitchB] traffic classifier c3

[SwitchB-classifier-c3] if-match vlan-id 100 to 200

[SwitchB-classifier-c3] quit

[SwitchB] traffic behaviorb3

[SwitchB-behavior-b3] nest top-most vlan-id 2

[SwitchB-behavior-b3] quit

[SwitchB] traffic classifier c4

[SwitchB-classifier-c4] if-match vlan-id 300 to 400

[SwitchB-classifier-c4] quit

[SwitchB] traffic behaviorb4

[SwitchB-behavior-b4] nest top-most vlan-id 3

[SwitchB-behavior-b4] quit

[SwitchB] traffic policy p2

[SwitchB-trafficpolicy-p2] classifier c3behaviorb3

[SwitchB-trafficpolicy-p2] classifier c4behaviorb4

[SwitchB-trafficpolicy-p2] quit

（3）在SwitchA和SwitchB的GE1/0/1端口上应用流策略实现灵活QinQ功能。

SwitchA上的配置：

[SwitchA] interfacegigabitethernet 1/0/1

[SwitchA-GigabitEthernet1/0/1] port link-type hybrid

[SwitchA-GigabitEthernet1/0/1] port hybrid Untagged vlan 2 3

[SwitchA-GigabitEthernet1/0/1] traffic-policy p1 inbound

[SwitchA-GigabitEthernet1/0/1] quit

SwitchB上的配置：

[SwitchB] interfacegigabitethernet 1/0/1

[SwitchB-GigabitEthernet1/0/1] port link-type hybrid

[SwitchB-GigabitEthernet1/0/1] port hybrid Untagged vlan 2 3

[SwitchB-GigabitEthernet1/0/1] traffic-policy p2 inbound

[SwitchB-GigabitEthernet1/0/1] quit

（4）配置SwitchA和SwitchB与运营商网络连接的GE1/0/2端口类型为Trunk或带标签的Hybird类型，并加入所需的外层VLAN中。

SwitchA上的配置：

[SwitchA] interface gigabitethernet 1/0/2

[SwitchA-GigabitEthernet1/0/2] port link-type trunk

[SwitchA-GigabitEthernet1/0/2] port trunk allow-pass vlan 2 3

[SwitchA-GigabitEthernet1/0/2] quit

SwitchB上的配置：

[SwitchB] interface gigabitethernet 1/0/2

[SwitchB-GigabitEthernet1/0/2] port link-type trunk

[SwitchB-GigabitEthernet1/0/2] port trunk allow-pass vlan 2 3

[SwitchB-GigabitEthernet1/0/2] quit

如果SwitchA、SwitchB上配置正确，则PC上网用户可以通过运营商网络互相通信；VoIP用户可以通过运营商网络互相通信。

7.6 QinQ映射配置与管理

QinQ映射功能可以将用户的VLAN标签映射为指定的VLAN标签，从而起到屏蔽不同用户VLAN标签的作用。但QinQ映射功能只能在子接口上应用，通过QinQ映射功能，子接口在向外发送本地VLAN的帧时，将帧中的本地VLAN标签替换成外部VLAN标签；在接收外部VLAN的帧时，又将帧中的外部VLAN标签替换成本地VLAN标签。QinQ 映射功能节省了大量的物理端口，因为一个物理端口可以划分许多子接口，而

每个子接口可以单独配置不同VLAN ID范围的标签映射。具体参见本章 7.3.4节介绍。

QinQ映射又分 1 to 1的映射方式和 2 to 2的映射方式，仅在**S5700**中的**S5710EI**、**S5700HI**、**S7700**、**S9300**和**S9700**系列中支持 。下面分别介绍其配置方法。

注意

物理端口和该端口下的子接口不能对同一VLAN进行VLAN映射或灵活QinQ配置。如果已经在子接口上配置QinQ映射功能，那么不能再在该子接口下配置灵活QinQ、QinQ终结、Dot1q终结相关命令。

7.6.1 配置1 to 1的QinQ映射

在子接口上部署1 to 1的QinQ映射功能后，当子接口收到带有一层标签的数据帧后，将数据帧中携带的一层标签映射为用户指定的一层标签。“1 to 1”的意思也可理解为直接进行标签替换，帧在映射前、后都只带有一层VLAN标签。

1 to 1的QinQ映射功能的配置方法很简单，只需要在对应子接口视图下使用 qinq mapping vidvlan-id1 [tovlan-id2] map-vlan vidvlan-id3命令即可。命令中的参数说明如下。

（1）vlan-id1 to vlan-id2：指定QinQ帧中原来一层标签的VLAN ID，其中 vlan-id1用来指定原标签的VLAN ID范围的起始值，取值范围均为 2～4 094的整数；可选参数vlan-id2用来指定原标签的VLAN ID范围的结束值，取值范围均为 3～4 094的整数，但vlan-id2必须大于vlan-id1。

（2）vlan-id3：指定映射后的一层标签的VLAN ID，取值范围为 1～4 094的整数。

缺省情况下，子接口下没有配置对报文中携带的标签进行映射操作，可用undo qinq mapping vid vlan-id1 [to vlan-id2] map-vlan vid vlan-id3命令取消子接口对应的QinQ映射功能。

注意

本命令用来配置子接口单层VLAN映射，且只对入方向报文生效。但子接口配置的转换前 VLAN 不能在全局下创建，也不能查看该 VLAN 信息。但映射前的标签和同一物理端口下的其他子接口下用来替换的外层标签互斥，即两者取值不能相同。

【示例 】配置Gigabitethernet0/0/1.1端口的QinQ映射功能，将外部VLAN 100替换为本地VLAN 200。

<HUAWEI>system-view

[HUAWEI] interface gigabitethernet 0/0/1.1

[HUAWEI-GigabitEthernet0/0/1.1] qinq mapping vid 100 map-vlan vid 200

7.6.2 配置2 to 1的QinQ映射

在子接口上部署2 to 1的QinQ映射功能，当子接口收到带有两层标签的数据帧后，将数据帧中携带的双层标签中的外层标签映射为用户指定的一层标签。即仅对帧中原来双层标签中的外层标签进行替换，内层标签当作数据部分使用，起到屏蔽内层标签的作用。

2 to 1的QinQ映射功能的配置方法也很简单，只需在对应的子接口视图下使用qinq mappingpe-vid vlan-id1 ce-vid vlan-id2 [to vlan-id3] map-vlan vid vlan-id4命令即可。命令中的参数说明如下。

（1）vlan-id1：指定帧中原来携带的外层标签的VLAN ID，取值范围是 2～4 094的整数。

（2）vlan-id2 [to vlan-id3]：指定帧中原来携带的内层标签的VLAN ID范围，其中vlan-id2用来指定内层标签的VLAN ID范围的起始值，取值范围为 1～4 094的整数；可选参数 vlan-id3用来指定内层标签的VLAN ID范围的结束值，取值范围为 2～4 094的整数，但vlan-id3必须大于vlan-id2。

（3）vlan-id4：指定映射后的外层标签的VLAN ID，取值范围为 1～4 094的整数。

缺省情况下，子接口下没有配置对帧中携带的标签进行映射操作，可用 undo qinq mapping pe-vid vlan-

id1 ce-vid vlan-id2 [to vlan-id3] map-vlan vid vlan-id4命令取消子接口替换带有双层标签的帧的外层标签。

注意

本命令用来配置子接口双层VLAN映射（只对外层VLAN映射，内层VLAN不 变，且只对入方向报文生效）。子接口配置的转换前 VLAN 不能在全局下创建，也不能查看该VLAN信息。且物理端口和该物理端口下的子接口不能对同一VLAN进行VLAN映射或灵活QinQ配置。

【示例 】将GE0/0/1.1子接口接收到的外层标签为10、内层标签为20的数据帧的外层标签替换为30。

```
<HUAWEI>system-view
[HUAWEI] interface gigabitethernet 0/0/1.1
[HUAWEI-GigabitEthernet0/0/1.1] qinq mapping pe-vid 10 ce-vid 20 map-vlan vid 30
```

7.7 VLAN映射基础

上节介绍了具有双层标签的QinQ VLAN数据帧的标签映射，本节要介绍的广泛意义上的帧（包括双层标签的QinQ帧、单层标签的VLAN帧和不带标签的普通帧）的标签映射，但这里均是在物理端口上应用的，不是像**QinQ**映射那样仅可应用于子接口上 。

虽然理论上可以有 4 096 个 VLAN 可用，但在实际的组网中，有些设备所支持的VLAN ID范围要远小于这个值，且还有一部分VLAN ID是保留不能用的。当用户网络有业务数据需要穿过 ISP网络时，就可能导致用户的VLAN ID与 ISP的公网VLAN ID相互冲突。为了解决这一问题，于是开发了本节将要介绍的VLAN映射（VLAN Mapping）技术。通过VLAN映射的配置可实现用户VLAN与运营商VLAN的相互映射，在数据帧到达配置了 VLAN 映射的交换机端口时替换用户数据帧中的 VLAN 标签，使用户业务按照运营商的VLAN ID规划进行传输。

7.7.1 VLAN映射原理

VLAN映射可以实现在用户VLAN ID和运营商VLAN ID之间相互转换。配置了VLAN映射功能后会在交换机内部维护一张VLAN映射表，然后对进入交换机的数据帧根据映射表进行VLAN映射操作。VLAN映射发生在数据帧从入交换机端口接收进来之后，到从出端口转发出去之前。交换机收到数据数据帧后，会根据帧中是否带有VLAN标签做出以下两种处理方式。

（1）数据帧带有 VLAN 标签：根据配置的 VLAN 映射方式，决定替换单层、双层或双层中的外层VLAN标签；然后进入MAC地址学习阶段，根据源MAC地址+映射后的VLAN ID刷新MAC地址表项；根据目的MAC+映射后VLAN ID查找MAC地址表项，如果没有找到，则在VLAN ID对应的VLAN内广播，否则从表项对应的端口转发。

（2）数据帧不带VLAN标签：根据配置的VLAN划分方式决定是否添加VLAN标签，对于不能加入VLAN 的数据帧上送 CPU 或丢弃，否则添加标签；然后进入 MAC 地址学习阶段，按照二层转发流程进行转发。

如图 7-19所示，如果在SwitchA的端口Port1上配置了VLAN 2和VLAN 3映射，则在端口Port1向外发送VLAN 2的帧时会将帧中的VLAN标签替换成VLAN 3；在接收VLAN 3的帧时又将帧中的VLAN标签替换成VLAN 2，然后按照二层转发流程进行数据转发，这样VLAN 2和VLAN 3就能实现互相通信。

当然，要想借助VLAN映射实现两个VLAN内设备互相通信，这两个VLAN内设备的IP地址还必须处于同一IP子网中，否则不同VLAN内设备间的互通需要依赖三层路由实现，此时就失去了VLAN映射的意义。

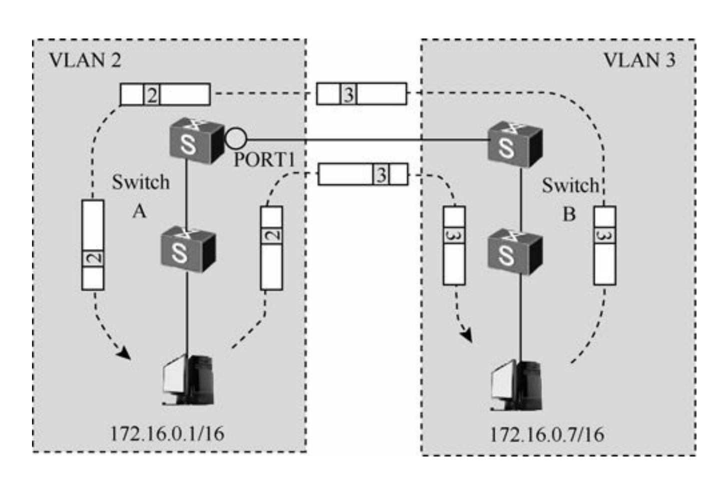


图7-19 VLAN映射应用示例

7.7.2 VLAN映射特性及产品支持

VLAN映射特性包括“VLAN映射方式”和“VLAN映射实现方式”两个方面。但是不同华为S交换机系列对VLAN映射特性的支持有较大区别。下面分别予以介绍。

1．VLAN映射实现方式

在VLAN映射实现方式方面，不同的交换机系列也有很大区别：在S2700和S3700系列交换机中仅支持基于VLAN的VLAN映射；在S5700和S6700系列交换机中仅支持基于 VLAN 的 VLAN 映射和基于流策略的VLAN 映射两种实现方式；而在高端的S7700、S9300和S9700系列中，支持基于VLAN的VLAN映射、基于802.1p优先级的VLAN 映射、基于 VLAN+802.1p 优先级组合方式的 VLAN 映射，以及基于流策略的VLAN映射共4种实现方式。以上总共4种VLAN映射实现方式的说明如表7-16所示。

表7-16 VLAN映射方式比较

VLAN 映射实现方式	映射原理
基于 VLAN	端口在接收到带有 VLAN 标签的数据帧后，依据帧中的 VLAN ID 进行映射操作，将帧中的 VLAN ID 映射为公网的 VLAN ID 基于 VLAN 的 VLAN 映射可以实现替换单层、双层或双层中的外层标签的功能
基于 802.1p 优先级	端口在接收到带有 VLAN 标签的数据帧后，依据帧中的 802.1p 优先级进行灵活的映射操作，将帧中的 VLAN ID 映射为公网的 VLAN ID 基于 802.1p 优先级的 VLAN 映射可以实现替换单层标签的功能

（续表）

VLAN 映射实现方式	映射原理
基于 VLAN+802.1p 优先级	端口在接收到带有 VLAN 标签的数据帧后，同时依据帧中的 802.1p 优先级和 VLAN ID 进行灵活的映射操作，将帧中的 VLAN ID 映射为公网的 VLAN ID 基于 VLAN+802.1p 优先级的 VLAN 映射可以实现替换单层标签的功能
基于流策略	通过配置流策略，对数据帧中的 VLAN ID 进行流分类，然后将流分类与某种流行为关联，对符合流分类的数据帧进行相应的处理（重标记数据帧的 VLAN ID 值），从而实现 VLAN 映射功能。基于流策略的 VLAN 映射能够针对业务类型提供差别服务 基于流策略的 VLAN 映射可以实现替换单层、双层或双层中的外层标签的功能

说明

当基于802.1p优先级的VLAN映射与基于VLAN的VLAN映射同时匹配时， 基于802.1p优先级的VLAN映射优先；当基于802.1p优先级的VLAN映射与基于VLAN的VLAN映射指定的映射前VLAN相同，但是基

于VLAN的VLAN映射指定了映射后的 802.1p优先级，而基于 802.1p优先级的 VLAN映射没有指定映射后的802.1p优先级时，映射后的802.1p优先级为基于VLAN的VLAN映射指定的优先级。

2．VLAN映射方式

目前，华为S系列交换机支持以下几种VLAN映射方式，当然不同交换机系列对不同映射方式的支持不同，下面具体介绍。

（1）1 to 1的映射方式

1 to 1的VLAN映射功能可使设备端口在接收到带有单层VLAN标签的数据帧时，将数据帧中携带的单层VLAN标签映射为公网的单层VLAN标签。这与前面7.6.1节介绍的 1 to 1的QinQ映射的功能是一样的但实现方式不一样。目前除了S1700和S2700SI系列外，其他华为S系列交换机均支持这种VLAN映射方式。

图 7-20所示为一个园区接入网中采用了1 to 1 VLAN映射方式。ISP端为每个家庭的不同业务采用了不同的VLAN（HSI、IPTV、VoIP分别对应VLAN 2、VLAN 3和VLAN 4）进行传输。

为了区分不同的家庭用户，需要在楼道交换机处将不同家庭用户的相同业务采用不同的VLAN进行发送，即进行1 to 1的VLAN映射。显然，这就需要提供大量的VLAN来隔离不同用户的不同业务，而汇聚层网络接入设备可以提供的 VLAN 数量有限（特别是在大的园区网络中），所以又需要在园区交换机上完成VLAN的汇聚功能，即将由多个VLAN发送的、不同用户的相同业务采用同一个VLAN进行发送，这其实又是“N to 1”的VLAN映射方式。

（2）2 to 1的映射方式

2 to 1的VLAN映射功能可使设备主接口在接收到带有双层VLAN标签的数据帧时，将数据帧中携带的外层标签映射为公网的标签，内层标签作为数据透传（也就是当做数据的一部分，不考虑内层的VLAN标签）。目前除了**S1700**、**S2700**、**S3700**和**S5700SI**系列外，其他华为S系列交换机均支持 **2 to 1**的**VLAN**映射功能。

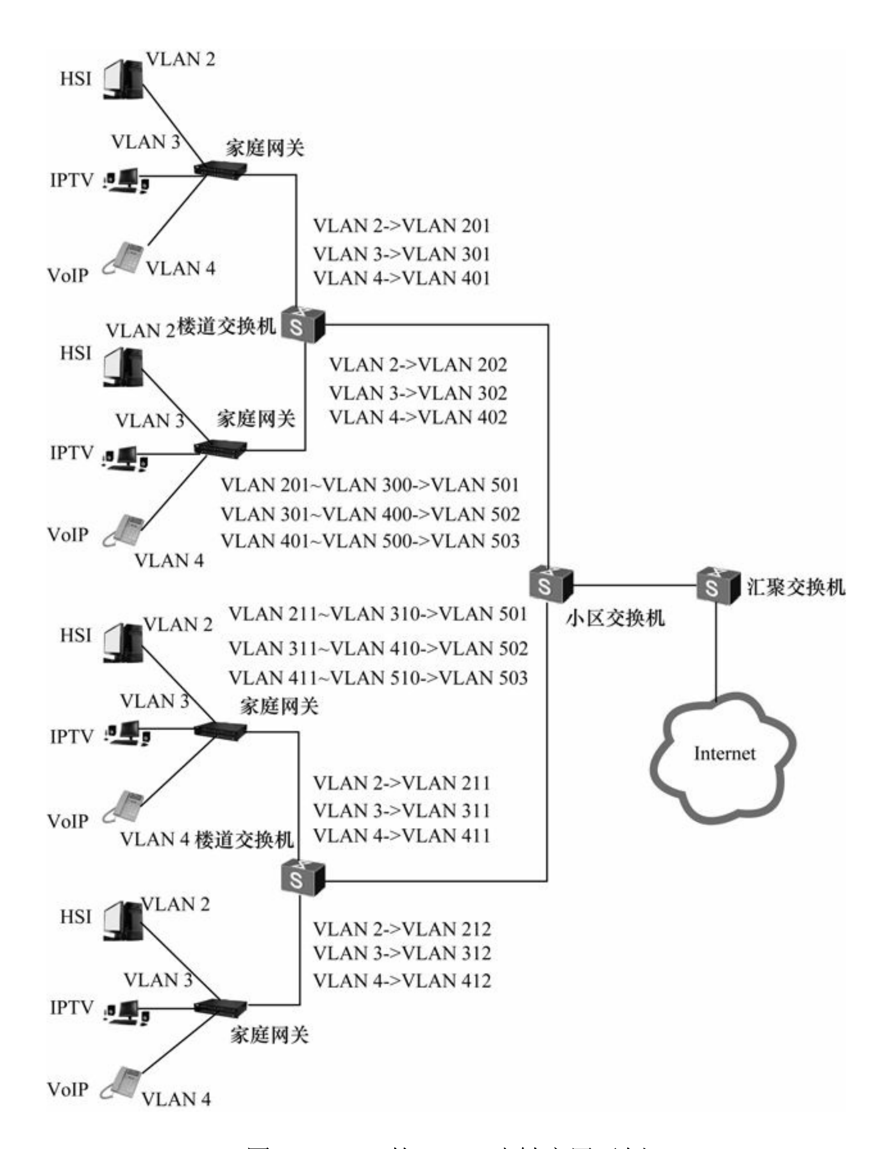


图7-20 1 to 1的VLAN映射应用示例

2 to 1的VLAN映射主要用于图 7-21所示的园区组网环境中。用户通过家庭网关、楼道交换机和小区交换机接入汇聚层网络。为了区分不同的用户和业务，以便进行网络管理和计费等，可以在楼道交换机上部署QinQ功能（具体将在本章后面介绍），以实现数据帧带有双层甚至多层 VLAN 标签。然后为了节约 VLAN 资源，在楼道交换机上配置N to 1的VLAN映射，以一个外层VLAN标签映射多个内层VLAN标签（如图中的VLAN 201标签映射VLAN 2和VLAN 3这两个VLAN），同时在小区交换机上部署 2 to 1 的 VLAN 映射功能，将不同用户的相同业务数据帧中的外层 VLAN 标签采用同一个VLAN标签进行替换（如用VALN 501标签替换原来VLAN 201标签和VLAN 401标签）。

（3）2 to 2的映射方式

2 to 2的VLAN映射功能可使设备主接口在接收到带有双层VLAN标签的数据帧时，将数据帧中携带的双层 VLAN 标签均映射为公网的双层 VLAN 标签。目前仅 S7700、S9300、S9300E和S9700高端系列交换机支持 2 to 2的VLAN映射功能。

2 to 2的VLAN映射主要用于图 7-22所示的组网环境。在本示例中，处于不同地理位置的用户为了可以规划自己的私网VLAN ID，不会导致和 ISP网络中的VLAN ID冲突，同时便于区分不同的用户和业务，采

用了 QinQ 方式传输，即用户数据帧中带有双层VLAN标签。但是由于用户数据帧中的VLAN ID与 ISP网络分配的VLAN ID不一致，将导致用户数据帧被丢弃，从而导致用户通信中断。这时可以在PE（提供商边缘）侧部署 2 to 2的VLAN映射功能，将用户网络的双层标签全部替换成 ISP网络的双层标签。如数据帧中的外层私网VLAN 100和VLAN 200均被替换成外层公网VLAN 50，内层私网VLAN 10和VLAN 20被替换成内层公网VLAN 60。

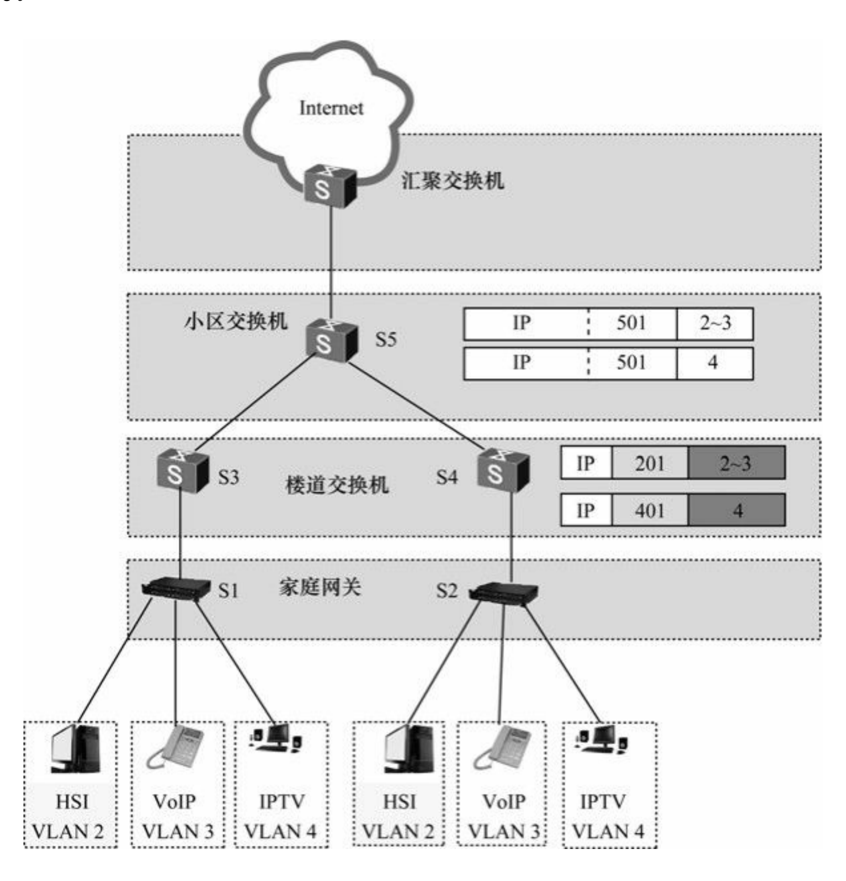


图7-21 2 to 1的VLAN映射应用示例

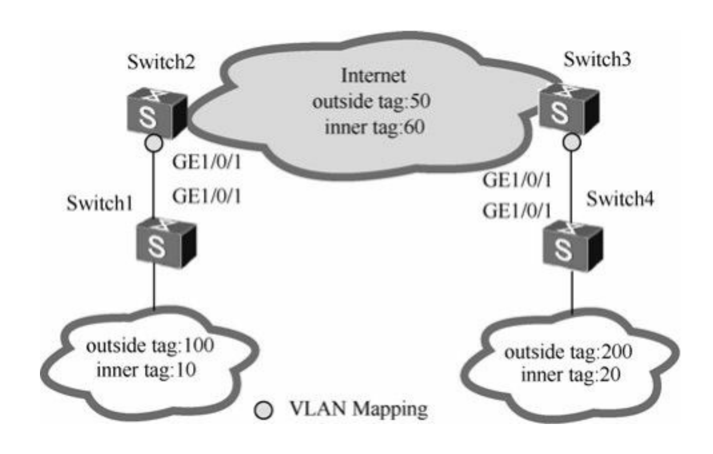


图7-22 2 to 2的VLAN映射应用示例

[**7.8 配置1 to 1的VLAN映射**]

1 to 1的VLAN映射功能可将数据帧中携带的单层VLAN标签映射为公网的单层VLAN标签。“1 to 1”的含义就是单层VLAN标签与单层VLAN标签之间的替换，这也就决定了在接入层交换机上无需使用生成双层甚至多层VLAN标签的QinQ协议。根据7.7.2节介绍的VLAN映射实现方式的不同，1 to 1的VLAN映射功能也有几种不同的配置方法，本节将分别予以介绍。

注意

1 to 1的VLAN映射功能必须在交换机与其他设备相连的Trunk或Hybrid类型端口上 配置。这些Trunk或Hybrid类型端口必须加入映射后的VLAN，但S系列交换机中的单板必须以带标签方式加入映射后的VLAN。在本节后面各小节介绍的1 to 1的VLAN映射功能配置中有关端口类型的允许VLAN的配置不再具体给出，在实际方案配置时一定要加上。

7.8.1 配置基于VLAN的1 to 1的VLAN映射

基于VLAN的VLAN映射功能可实现端口在接收到带有单层VLAN标签的数据帧后，依据帧中的VLAN ID进行映射操作，将帧中的VLAN ID映射为指定的VLAN ID。其配置方法很简单，只需要指定VLAN映射中的源VLAN和目的VLAN（当然这些VLAN必须得事先已创建好），然后在对应的交换机端口上启用VLAN 映射功能即可。具体如表7-17所示。

表7-17 基于VLAN的1 to 1的VLAN映射的配置步骤

步骤	命令	说明
1	**system-view** 例如：< HUAWEI > **system-view**	进入系统视图
2	**interface** *interface-type interface-number* 例如：[HUAWEI]**interface gigabitethernet 0/0/1**	键入要配置 1 to 1 VLAN 映射的交换机端口，进入接口视图
3	**qinq vlan-translation enable** 例如：[HUAWEI-GigabitEthernet0/0/1] **qinq vlan-translation enable**	使能端口的 VLAN 转换功能。只有在端口使能了 VLAN 转换功能后，才可以在端口上配置 VLAN 映射和灵活 QinQ 功能。**本命令仅 S2700、S3700、S5700 和 S6700 系列需要配置** 缺省情况下，没有使能端口的 VLAN 转换功能，可用 **undo qinq vlan-translation enable** 命令取消端口的 VLAN 转换功能
4	**port vlan-mapping vlan** *vlan-id1* [**to** *vlan-id2*] **map-vlan** *vlan-id3* [**remark-8021p** *8021p-value*] 例如：[HUAWEI-GigabitEthernet0/0/1]**port** GigabitEthernet0/0/1 **to** 0/0/4	配置端口的单层标签的 VLAN 映射功能。命令中的参数说明如下。 （1）*vlan-id1* [**to** *vlan-id2*]：指定要映射的源 VLAN ID，其中 *vlan-id1* 表示起始 VLAN ID，可选参数 **to** *vlan-id2* 表示结束 VLAN ID，取值范围均为 1～4 094 的整数。同一个交换机端口推荐最多指定 16 个映射前 VLAN （2）*vlan-id3*：指定映射后的 VLAN ID，取值范围也为 1～4 094 的整数

（续表）

步骤	命令	说明
4	**port vlan-mapping vlan** *vlan-id1* [**to** *vlan-id2*] **map-vlan** *vlan-id3* [**remark-8021p** *8021p-value*] 例如：[HUAWEI-GigabitEthernet0/0/1]**port** GigabitEthernet0/0/1 **to** 0/0/4	（3）*8021p-value*：可选参数，指定映射后的 VLAN 帧的 802.1p 优先级。802.1p 是 802.1Q 的 VLAN 帧中的 PRI（Priority）字段，长度为 3bit，用于当交换设备阻塞时，优先发送优先级高的数据包。通过本可选参数的设置，可实现端口在接收到带 VLAN 标签的数据帧后，将帧中的 802.1p 优先级修改为用户配置的 802.1p 优先级值 缺省情况下，交换机端口下没有配置对数据帧中携带的 VLAN 标签进行映射操作，可用 **undo port vlan-mapping** { **all** | **vlan** *vlan-id1* [**to** *vlan-id2*] [**map-vlan** *vlan-id3*]} 命令取消对数据帧中携带的指定或所有单层 VLAN 标签进行映射操作

注意

当映射前的源VLAN的取值为vlan-id1 to vlan-id2指定的一个范围VLAN时，则 该交换机端口需要以带标签的方式加入这些源VLAN，并且参数vlan-id3所对应映射后的VLAN不允许配置成VLANIF接口。如果同时配置了VLAN映射和DHCP功能，则端口需要以带标签的方式加入映射前的这些VLAN，因为DHCP服务器要依据映射前的VLANIF接口来为不同VLAN分配不同的IP地址。

【示例 】将GE0/0/1端口接收到的VLAN ID为 100的数据帧映射为VLAN ID为 10的数据帧，然后进行转发。

```
<HUAWEI>system-view
[HUAWEI] interface gigabitethernet0/0/1
[HUAWEI-GigabitEthernet0/0/1] port link-type trunk
[HUAWEI-GigabitEthernet0/0/1] port trunk allow-pass vlan 10
[HUAWEI-GigabitEthernet0/0/1] qinq vlan-translation enable
[HUAWEI-GigabitEthernet0/0/1] port vlan-mapping vlan 100 map-vlan 10
```

7.8.2 配置基于802.1p优先级的1 to 1的VLAN映射

配置基于802.1p优先级的VLAN映射功能，可以根据进入端口的数据帧的802.1p优先级和 VLAN ID进行灵活的映射，优先保证重要用户的正常通信。但基于 802.1p优先级的VLAN映射功能仅在S7700、S9300和S9700系列E系列和F系列单板支持。

基于 802.1p优先级的 1 to 1的VLAN映射是在入端口上配置的，但如果在入端口上创建了DiffServ域并配置优先级映射关系，此时内部优先级和802.1p优先级可能不一样，这时还需要在出端口上配置VLAN优先级映射。这与7.5.3节介绍的基于802.1p优先级的灵活QinQ功能类似。

在入端口上配置基于802.1p优先级的VLAN映射的具体配置步骤如表7-18所示。注意，这里面包括单独采用基于802.1p优先级映射实现方式，和同时采用基于VLAN+802.1p优先级映射实现方式下的1 to 1的VLAN映射配置，根据实际需要选择一种实现方式即可。

表7-18 在入端口上配置基于802.1p优先级的VLAN映射的配置步骤

步骤	命令	说明
1	**system-view** 例如：<HUAWEI> **system-view**	进入系统视图

（续表）

步骤	命令	说明
2	**Interface** *interface-type interface-number* 例如：[HUAWEI]**interface gigabitethernet 0/0/1**	键入要配置 1 to 1 VLAN 映射的交换机端口，进入接口视图
3	**port vlan-mapping 8021p** *8021p-value* **map-vlan** *vlan-id* [**remark-8021p** *8021p-value2*] 例如：[HUAWEI-GigabitEthernet0/0/1]**port vlan-mapping 8021p** 5 **map-vlan** 200	（二选一）配置以上入端口基于 802.1p 优先级的 VLAN 映射功能。命令中的参数说明如下。 （1）*8021p-value1*：指定外层 VLAN 的 802.1p 优先级，取值范围为 0～7 的整数，值越大优先级越高 （2）*vlan-id*：指定映射后的 VLAN ID，取值范围为 1～4 094 的整数 （3）*8021p-value2*：可选参数，指定使用映射的 VLAN 标签进行修改后帧的 802.1p 优先级。通过本可选参数的设置，可实现端口在接收到带 VLAN 标签的数据帧后，将帧中的 802.1p 优先级修改为用户配置的 802.1p 优先级值 缺省情况下，交换机端口下没有配置基于 802.1p 优先级的 VLAN 映射功能，可用 **undo port vlan-mapping 8021p** *8021p-value1* [**map-vlan** *vlan-id*] 命令取消对应交换机端口的基于 802.1p 优先级的 VLAN 映射功能
	port vlan-mapping vlan *vlan-id1* [**to** *vlan-id2*] **8021p** *8021p-value1* [**to** *8021p-value2*] **map-vlan** *vlan-id3* [**remark-8021p** *8021p-value3*] 例如：[HUAWEI-GigabitEthernet0/0/1]**port vlan-mapping vlan 100 8021p 5 map-vlan 200 remark-8021p** 1	（二选一）配置以上入端口基于 VLAN+802.1p 优先级的 VLAN 映射功能。命令中的参数说明如下。 （1）*vlan-id1* [**to** *vlan-id2*]：指定要映射的源 VLAN ID 的取值范围，其中 *vlan-id1* 表示起始 VLAN ID，可选参数 **to** *vlan-id2* 表示结束 VLAN ID，取值范围均为 1～4094 的整数。同一个交换机端口推荐最多指定 16 个映射前 VLAN （2）*8021p-value1* [**to** *8021p-value2*]：指定要映射的外层 VLAN 的 802.1p 优先级的取值范围，其中 *8021p-value1* 表示 802.1p 优先级取值范围的下限，取值范围为 0～7 的整数，值越大优先级越高；**to** *8021p-value2* 表示 802.1p 优先级取值范围的上限，取值范围为 1～7 的整数，*8021p-value2* 的取值必须大于 *8021p-value1* 的取值 （3）*vlan-id3*：指定映射后的 VLAN ID，取值范围为 1～4 094 的整数 （4）*8021p-value3*：可选参数，指定映射后的 VLAN 的 802.1p 优先级，取值范围为 0～7 的整数，值越大优先级越高。通过本可选参数的设置，可实现端口在接收到带 VLAN 标签的数据帧后，将帧中的 802.1p 优先级修改为用户配置的 802.1p 优先级值 缺省情况下，交换机端口下没有配置对数据帧中携带的 VLAN 标签进行映射操作，可用 **undo port vlan-mapping vlan** *vlan-id1* [**to** *vlan-id2*] [**8021p** *8021p-value1* [**to** *8021p-value2*]] [**map-vlan** *vlan-id3*]命令取消对应交换机端口基于 VLAN+802.1p 优先级的 VLAN 映射功能

【示例 1】将进入GE1/0/1端口的802.1p优先级为5的帧映射为VLAN ID为200的帧。

<HUAWEI>system-view

[HUAWEI] interface gigabitethernet1/0/1

[HUAWEI-GigabitEthernet1/0/1] port vlan-mapping 8021p 5 map-vlan 200

【示例 2】将进入GE1/0/1端口的VLAN ID为 100、802.1p优先级为 5的帧映射为VLAN ID为 200的帧，并修改映射后的 802.1p优先级为 1，再进行转发。

<HUAWEI>system-view

[HUAWEI] interface gigabitethernet1/0/1

[HUAWEI-GigabitEthernet1/0/1] port vlan-mapping vlan 100 8021p 5 map-vlan 200 remark-8021p 1

如果在入端口创建了 DiffServ（差分服务）域，并配置了优先级映射关系，此时数 据帧中的内部优先级和 802.1p 优先级可能不一样，此时除了要进行表 7-20 所示的配置外，还需在出端口上配置优先级映射关系。具体配置步骤参见表7-16。

7.8.3 配置基于流策略的1 to 1的VLAN映射

这里所说的“流策略”是指将流分类和流行为关联后形成的完整的QoS策略。QoS策略的配置包括以下4个基本任务：（1）定义流分类；（2）定义流行为；（3）创建QoS策略，将流分类与某种流行为进行关联；（4）在端口上应用QoS策略。

根据以上4个QoS策略的基本配置任务可以得出本节的VLAN映射基本配置任务如下。

（1）根据数据帧中的VLAN ID进行流分类。

（2）定义基于VLAN ID的流行为。

（3）创建QoS策略，将以上定义的流分类与流行为进行关联，对符合流分类的数据帧进行相应的处理（重标记数据帧的VLAN ID值），从而实现VLAN映射功能。

（4）在交换机端口上应用以上QoS策略，以针对业务类型提供差别服务。

基于流策略的 1 to 1的 VLAN策略可以在端口的入方向上应用，对接收的帧进行VLAN标签替换，或者在端口的出方向上应用，对发出的帧进行VLAN标签替换。但两种配置方法类似，具体配置步骤如表7-19所示。它与7.5.4节介绍的基于流策略的灵活QinQ配置步骤类似，但灵活**QinQ**仅可在入方向上应用 。

表7-19 基于流策略的1 to 1的VLAN映射的配置步骤

配置任务	步骤	命令	说明
定义流分类	1	**system-view** 例如：< HUAWEI > **system-view**	进入系统视图
	2	**traffic classifier** *classifier-name* 例如：[HUAWEI]**traffic classifier** c1	创建流分类并进入流分类视图。其他说明参见表7-15 中的第 2 步
	3	**if-match vlan-id** *start-vlan-id* [**to** *end-vlan-id*] 例如：[HUAWEI-classifier-c1]**if-match vlan-id** 2	用来在流分类中创建基于 VLAN ID 进行分类的匹配规则。其他说明参见表 7-15 中的第 3 步
	4	**quit** 例如：[HUAWEI-classifier-c1]**quit**	退出流分类视图，返回系统视图
定义流行为	5	**traffic behavior** *behavior-name* 例如：[HUAWEI] **traffic behavior** b1	创建流行为并进入流行为视图。其他说明参见表7-15 中的第 5 步
	6	**remark vlan-id** *vlan-id* 例如：[HUAWEI-behavior-b1]**remark vlan-id** 200	配置流行为，在流行为中创建重标记 VLAN 数据帧的外层 VLAN 标签值。参数 *vlan-id* 用来指定重标记数据帧的 VLAN ID 值（也即映射后的 VLAN ID），取值范围为 1～4 094 的整数 缺省情况下，流行为中没有重标记 VLAN 数据帧的标签值的动作，可用 **undo remark vlan-id** 命令恢复缺省情况
	7	**quit** 例如：[HUAWEI-behavior-b1] **quit**	退出流行为视图，返回系统视图

（续表）

配置任务	步骤	命令	说明
创建 QoS 策略，关联流分类与流行为	8	**traffic policy** *policy-name* 例如：[HUAWEI] **traffic policy** p1	创建流策略并进入流策略视图。其他说明参见表 7-15 中的第 8 步
	9	**Classifier** *classifier-name* **behavior** *behavior-name* 例如：[HUAWEI-trafficpolicy-p1]**classifier** c1 **behavior** b1	将以上定义的流分类与指定的流行为进行绑定，组成流策略。其他说明参见表 7-15 中的第 9 步
	10	**quit** 例如：[HUAWEI-dsdomain-ds1] **quit**	退出流策略视图，返回系统视图
在交换机端口的入出方向上应用 QoS 策略	11	**Interface** *interface-type interface-number* [HUAWEI] **interface gigabitethernet 1/0/1**	键入要应用流策略的交换机端口，进入接口视图
	12	**port link-type hybrid** 例如：[HUAWEI-GigabitEthernet1/0/1] **port link-type hybrid**	配置以上端口的类型为 Hybrid
	13	**port hybrid Untagged vlan** { { *vlan-id1* [**to** *vlan-id2*] }&<1-10> | **all** } 例如：[HUAWEI-GigabitEthernet1/0/1] **port hybrid Untagged vlan** 2 3	把以上 Hybrid 端口以不带标签方式加入到映射后的外层 VLAN 中，具体命令参数本第 6 章已有详细介绍，不再赘述
	14	**traffic-policy** *policy-name* { **inbound** | **outbound** } 例如：[HUAWEI-GigabitEthernet1/0/1] **traffic-policy p1 inbound**	在端口的入或出方向应用流策略。命令中的参数和选项说明如下。 *policy-name*：指定要应用的 QoS 流策略名 （1）**inbound**：二选一选项，指定在端口的入方向上应用指定的流策略 （2）**outbound**：二选一选项，指定在端口的出方向上应用指定的流策略 缺省情况下，交换机端口上没有应用任何流策略，可用 **undo traffic-policy** [*policy-name*] { **inbound** | **outbound** }命令取消在端口上应用流策略

【示例 】在新创建的流策略p1中，配置流分类c1关联流行为b1，并在GE1/0/1端口的入方向上应用该流策略。

<HUAWEI>system-view

[HUAWEI] traffic policy p1

[HUAWEI-trafficpolicy-p1] classifier c1 behavior b1

[HUAWEI-trafficpolicy-p1] quit

[HUAWEI] interface gigabitethernet 1/0/1

[HUAWEI-GigabitEthernet1/0/1] traffic-policy p1 inbound

7.8.4 基于VLAN的1 to 1VLAN映射配置示例

本示例拓扑结构如图7-23所示。不同的小区拥有相同的业务，如上网、IPTV、VoIP 等业务。为了便于管理，各个小区的网络管理者将不同的业务划分到不同的VLAN中，相同的业务划分到同一个VLAN中。但是，由于各小区网络管理者事先并没有协商好，目前存在不同的小区中相同的业务所属的VLAN不相同，但又需要实现相同业务、不同VLAN间的用户相互通信。如小区1和小区2中拥有相同的业务，但是属于不同的VLAN （即VLAN 5和VLAN 6，但这两个VLAN中的用户计算机同处一个 IP网段中）。现需要通过VLAN映射功能实现小区1和小区2中的用户可以直接互通。

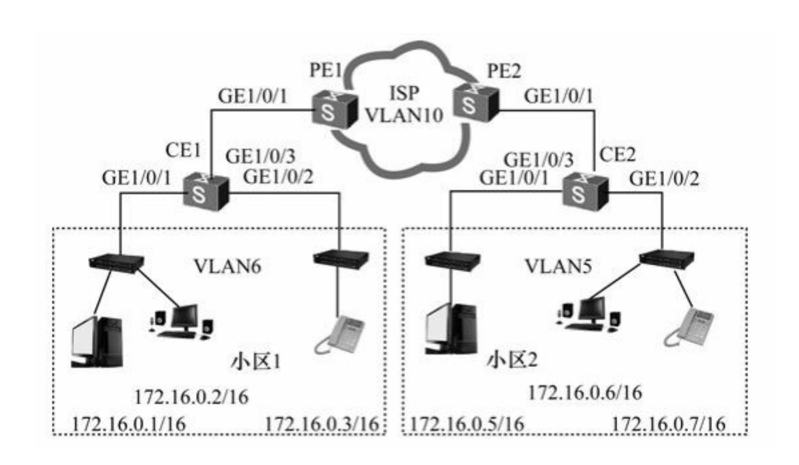


图7-23 基于VLAN的1 to 1的VLAN映射配置示例拓扑结构

1．配置思路分析

根据本示例的实际网络环境和应用需求，可以得出如下基本配置思路。

（1）将连接小区1中某业务用户连接的交换机端口以基于端口划分方式加入到VLAN 6中，将连接小区2中相同业务用户连接的交换机端口以基于端口划分方式加入VLAN 5中，用来区分不同的用户。这方面的配置在此不作介绍，具体可参见本书第6章6.2节。

（2）在运营商网络的边缘设备PE1和PE2的GE1/0/1端口上配置VLAN映射功能，将两个小区中的用户VLAN ID映射为运营商提供的同一个VLAN ID（VLAN 10），以实现原来两小区中不同VLAN用户间的直接互通。

2．配置方法

在此仅介绍以上第二项配置任务，即直接介绍VLAN映射的配置方法。具体配置步骤如下。

PE1上的配置：

```
<HUAWEI>system-view
[HUAWEI] sysname PE1
[PE1] vlan 10
[PE1-vlan10] quit
[PE1] interfacegigabitethernet 1/0/1
[PE1-GigabitEthernet1/0/1] port link-type trunk
[PE1-GigabitEthernet1/0/1] port trunk allow-pass vlan 10
[PE1-GigabitEthernet1/0/1] port vlan-mapping vlan 6 map-vlan 10
[PE1-GigabitEthernet1/0/1] quit
```

PE2上的配置：

```
<HUAWEI>system-view
[HUAWEI] sysname PE2
[PE2] vlan 10
[PE2-vlan10] quit
[PE2] interfacegigabitethernet1/0/1
[PE2-GigabitEthernet1/0/1] port link-type trunk
[PE2-GigabitEthernet1/0/1] port trunk allow-pass vlan 10
```

[PE2-GigabitEthernet1/0/1] port vlan-mapping vlan 5 map-vlan 10

[PE2-GigabitEthernet1/0/1] quit

配置好后，小区1中的用户和小区2中的用户可以互相Ping得通，表明配置成功。

7.9 配置2 to 1的VLAN映射

当网络边缘设备收到的数据帧带有两层VLAN标签时，内层标签代表用户，外层标签代表业务。为了区分不同的业务进入运营商网络，可在网络边缘设备上配置 2 to 1的VLAN映射功能。将代表不同业务的外层VLAN标签映射为用户指定的VLAN标签，内层标签作为数据透传到运营商网络，实现不同用户之间的通信。即这里的“2 to 1”的含义是对帧中原来两层VLAN标签中的外层标签进行替换。

本节介绍的 2 to 1的VLAN映射与 1 to 1的VLAN映射的配置总体上差不多，只是一些主要命令有所区别。而且，配置**2 to 1**的**VLAN**映射功能的端口类型也必须为**Trunk**或**Hybrid**， 且端口必须加入映射后的VLAN，S系列交换机中的单板也必须以带标签的方式加入映射后的VLAN。

7.9.1 配置基于VLAN的2 to 1的VLAN映射

基于VLAN的 2 to 1的VLAN映射功能可实现端口在接收到带有VLAN标签的帧后，依据帧中的内层VLAN ID进行外层标签映射操作，将帧中的外层标签映射为指定的标签。但要注意，一定要先在接入层交换机的对应端口上启用 QinQ 协议，以在帧中生成双层VLAN标签。

基于VLAN的 2 to 1的VLAN映射的具体配置方法很简单，具体如表 7-20所示（与7.8.1节介绍的 1 to 1的VLAN映射配置步骤差不多）。

表7-20 基于VLAN的2 to 1的VLAN映射的配置步骤

步骤	命令	说明
1	**system-view** 例如：< HUAWEI > **system-view**	进入系统视图
2	**Interface** *interface-type interface-number* 例如：[HUAWEI]**interface gigabitethernet 0/0/1**	键入要配置 1 to 1 VLAN 映射的交换机端口，进入接口视图
3	**qinq vlan-translation enable** 例如：[HUAWEI-GigabitEthernet0/0/1] **qinq vlan-translation enable**	使能端口的 VLAN 转换功能。只有在端口使能了 VLAN 转换功能后，才可以在端口上配置 VLAN 映射和灵活 QinQ 功能。**本命令仅 S2700、S3700、S5700 和 S6700 系列需要配置** 缺省情况下，没有使能端口的 VLAN 转换功能，可用 **undo qinq vlan-translation enable** 命令取消端口的 VLAN 转换功能

（续表）

步骤	命令	说明
4	**port vlan-mapping vlan** *vlan-id1* **inner-vlan** *vlan-id2* [**to** *vlan-id3*] **map-vlan** *vlan-id4* [**remark-8021p** *8021p-value*] 例如：[HUAWEI-GigabitEthernet0/0/1] **port vlan-mapping vlan** 10 **inner-vlan** 20 **map-vlan** 100	替换帧中的双层 VLAN 标签中的外层标签即可。执行本命令可实现端口在接收到帧后，依据帧中的内层标签 VLAN ID 进行外层标签的映射操作，将帧中的外层标签映射为指定的标签。命令中的各参数说明如下。 （1）*vlan-id1*：指定要映射的源外层标签的 VLAN ID，取值范围是 1～4 094 的整数 （2）*vlan-id2* [**to** *vlan-id3*]：指定要映射的源内层标签的 VLAN ID 范围，其中 *vlan-id2* 用来指定端口接收到的帧携带的内层标签的 VLAN ID 范围的起始值，可选参数 *vlan-id3* 用来指定端口接收到的帧携带的内层标签的 VLAN ID 范围的结束值，取值范围均为 1～4 094 的整数 （3）*vlan-id4*：指定帧中映射后的外层标签的 VLAN ID，取值范围是 1～4 094 的整数 （4）*8021p-value*：可选参数，指定修改映射后的 VALN 标签的 802.1p 优先级，取值范围为 0～7，值越大优先级越高。选择此可选参数可将帧中的 802.1p 优先级修改为用户配置的 802.1p 优先级值 缺省情况下，交换机端口下没有配置对数据帧中携带的 VLAN 标签进行映射操作，可用 **undo port vlan-mapping** { **all** | **vlan** *vlan-id1* **inner-vlan** *vlan-id2* [**to** *vlan-id3*] [**map-vlan** *vlan-id4*] }取消替换带有指定双层 Tag 的帧的外层标签 VLAN 映射操作

注意

如果在交换机上同时配置了VLAN映射和DHCP服务器功能，接口还需要以带标签的方式加入映射前VLAN。当数据帧同时匹配所配置的单、双层VLAN映射时，以精确匹配生效，即双层VLAN映射生效。

【示例 】配置2 to 1的VLAN映射功能，将内层VLAN标签为 20的帧中的外层标签10映射为外层标签100，然后进行转发。

```
<HUAWEI>system-view
[HUAWEI] interface gigabitethernet0/0/1
[HUAWEI-GigabitEthernet0/0/1] port link-type trunk
[HUAWEI-GigabitEthernet0/0/1] port trunk allow-pass vlan 100
[HUAWEI-GigabitEthernet0/0/1] qinq vlan-translation enable
[HUAWEI-GigabitEthernet0/0/1] port vlan-mapping vlan 10 inner-vlan 20 map-vlan 100
```

7.9.2 配置基于流策略的2 to 1的VLAN映射

与 7.8节介绍的 1 to 1的VLAN映射配置一样，2 to 1的VLAN映射也可以通过流策略进行。用户可以根据数据帧中外层标签的VLAN ID对流进行分类，然后将流分类与所定义的某种流行为关联，对符合流分类的数据帧进行相应的处理（重标记数据帧的VLAN ID值），从而实现VLAN映射功能。

在具体配置上，基于流策略的 2 to 1的VLAN映射也与 7.4.3节介绍的基于流策略的 1 to 1的 VLAN映射的配置方法和配置命令都基本一样。所配置的基本任务也就是QoS策略的四大配置任务：（1）定义流分类；（2）定义流行为；（3）创建QoS策略，将流分类与某种流行为进行关联；（4）在端口上应用QoS策略。可在交换机端口的入方向 （应用于接收的帧）或出方向（应用于发出的帧）上进行 2 to 1的VLAN映射配置，具体步骤如表7-21所示。

表7-21 基于流策略的2 to 1的VLAN映射的配置步骤

配置任务	步骤	命令	说明
定义流分类	1	**system-view** 例如：< HUAWEI > **system-view**	进入系统视图
	2	**traffic classifier** *classifier-name* **operator and** 例如：[HUAWEI]**traffic classifier** c1	创建流分类并进入流分类视图，参数 *classifier-name* 用来指定流分类名称，并指定流分类下各规则之间是逻辑“与”的关系。指定该逻辑关系后，当流分类中有 ACL 规则时，数据帧必须匹配其中一条 ACL 规则以及所有非 ACL 规则才属于该类；当流分类中没有 ACL 规则时，则数据帧必须匹配所有非 ACL 规则才属于该类 缺省情况下，系统没有定义任何流分类，可用 **traffic classifier** *classifier-name* 命令删除指定的流分类
	3	**if-match vlan-id** *vlan-id* 例如：[HUAWEI-classifier-c1]**if-match vlan-id** 2	配置匹配数据帧的规则，即指定数据帧的源外层 VLAN ID，参数 *vlan-id* 用来指定匹配的源外层 VLAN ID，取值范围为 1～4 094 的整数 缺省情况下，没有基于 VLAN ID 分类的匹配规则，可用 **undo if-match vlan-id** *vlan-id1* 删除指定的匹配规则
	4	**if-match cvlan-id** *cvlan-id* 例如：[HUAWEI-classifier-c1] **if-match cvlan-id** 20	配置匹配数据帧的规则，即指定数据帧的源内层 VLAN ID。参数 *cvlan-id* 用来指定匹配的源内层 VLAN ID，取值范围为 1～4 094 的整数 缺省情况下，流分类中没有基于 QinQ 数据帧内外两层 VLAN ID 进行分类的匹配规则，可用 **undo if-match cvlan-id** *vlan-id2* 命令在流分类中删除指定的匹配规则
	5	**quit** 例如：[HUAWEI-classifier-c1]**quit**	同流分类视图，返回系统视图
定义流行为	6	**traffic behavior** *behavior-name* 例如：[HUAWEI] **traffic behavior** b1	创建流行为并进入流行为视图。其他说明参见表 7-15 的第 5 步
	7	**remark vlan-id** *vlan-id* 例如：[HUAWEI-behavior-b1]**remark vlan-id** 200	配置流行为，在流行为中创建重标记帧中的外层标签 VLAN ID。其他说明参见表 7-19 中的第 6 步
	8	**quit** 例如：[HUAWEI-behavior-b1] **quit**	退出流行为视图，返回系统视图
创建 QoS 策略，关联流分类与流行为	9	**traffic policy** *policy-name* 例如：[HUAWEI] **traffic policy** p1	创建流策略并进入流策略视图。其他说明参见表 7-15 的第 8 步

（续表）

配置任务	步骤	命令	说明
创建 QoS 策略，关联流分类与流行为	10	**classifier** *classifier-name* **behavior** *behavior-name* 例如：[HUAWEI-trafficpolicy-p1]**classifier** c1 **behavior** b1	将以上定义的流分类与指定的流行为进行绑定，组成流策略。其他说明参见表 7-15 的第 9 步
	11	**quit** 例如：[HUAWEI-dsdomain-ds1] **quit**	退出流策略视图，返回系统视图
在交换机端口的入或出方向上应用 QoS 策略	12	**interface** *interface-type interface-number* [HUAWEI] **interface gigabitethernet 1/0/1**	键入要应用流策略的交换机端口，进入接口视图
	13	**port link-type hybrid** 例如：[HUAWEI-GigabitEthernet1/0/1] **port link-type hybrid**	配置以上端口的类型为 Hybrid
	14	**port hybrid Untagged vlan** { { *vlan-id1* [**to** *vlan-id2*] }&<1-10> | **all** } 例如：[HUAWEI-GigabitEthernet1/0/1] **port hybrid Untagged vlan** 2 3	把以上 Hybrid 端口以不带标签方式加入到映射后的外层 VLAN 中，具体命令参数本第 6 章已有详细介绍，不再赘述
	15	**traffic-policy** *policy-name* { **inbound** | **outbound** } 例如：[HUAWEI-GigabitEthernet1/0/1] **traffic-policy p1 inbound**	在端口的入或出方向应用流策略。其他说明参见表 7-19 的第 14 步

7.9.3 基于VLAN的2 to 1的VLAN映射配置示例

本示例拓扑结构如图 7-24 所示，用户通过家庭网关、楼道交换机和小区交换机接入汇聚层网络。为了节省运营商网络 VLAN 资源，及实现不同用户相同业务在传输过程中相互隔离，可以在楼道交换机上部署QinQ功能，在小区交换机上部署VLAN映射功能。

1．配置思路分析

根据本示例的要求，可采用如下的思路配置 2 to 1的VLAN映射，以实现VLAN资源节约。

（1）将连接用户的交换机端口分别划分到指定VLAN中，以区分不同的业务。

（2）在楼道交换机上部署QinQ功能，在帧中实现双层VLAN标签，以区分用户、业务。

（3）在小区交换机上部署VLAN映射功能，节约VLAN资源。

2．具体配置步骤

S1上的配置 ：将S1的下行口划分到指定的业务VLAN中。

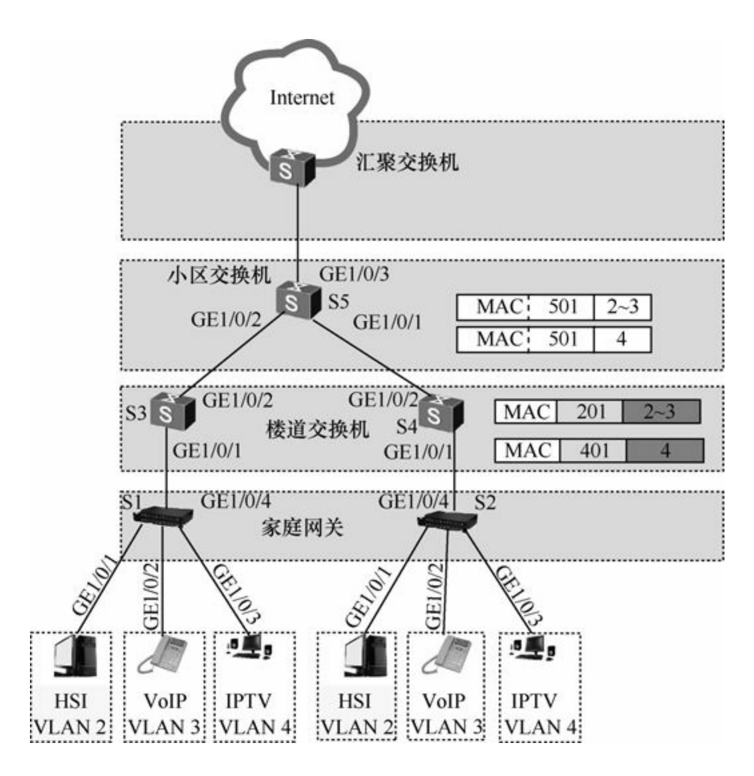


图7-24 基于VLAN的2 to 1的VLAN映射配置示例的拓扑结构

```
<HUAWEI>system-view
[HUAWEI] sysname S1
[S1] vlan batch 2 to 4
[S1] interfacegigabitethernet 1/0/1
[S1-GigabitEthernet1/0/1] port link-type access
[S1-GigabitEthernet1/0/1] port default vlan 2
[S1-GigabitEthernet1/0/1] quit
[S1] interface gigabitethernet 1/0/2
[S1-GigabitEthernet1/0/2] port link-type access
[S1-GigabitEthernet1/0/2] port default vlan 3
[S1-GigabitEthernet1/0/2] quit
[S1] interface gigabitethernet 1/0/3
[S1-GigabitEthernet1/0/3] port link-type access
[S1-GigabitEthernet1/0/3] port default vlan 4
[S1-GigabitEthernet1/0/3] quit
[S1] interface gigabitethernet 1/0/4
[S1-GigabitEthernet1/0/4] port link-type trunk
[S1-GigabitEthernet1/0/4] port trunk allow-pass vlan 2 to 4
[S1-GigabitEthernet1/0/4] quit
```

S2上的配置：将S2的下行口划分到指定的业务VLAN中。

```
<HUAWEI>system-view
```

[HUAWEI] sysname S2
[S2] vlan batch 2 to 4
[S2] interface gigabitethernet 1/0/1
[S2-GigabitEthernet1/0/1] port link-type access
[S2-GigabitEthernet1/0/1] port default vlan 2
[S2-GigabitEthernet1/0/1] quit
[S2] interfacegigabitethernet 1/0/2
[S2-GigabitEthernet1/0/2] port link-type access
[S2-GigabitEthernet1/0/2] port default vlan 3
[S2-GigabitEthernet1/0/2] quit
[S2] interfacegigabitethernet 1/0/3
[S2-GigabitEthernet1/0/3] port link-type access
[S2-GigabitEthernet1/0/3] port default vlan 4
[S2-GigabitEthernet1/0/3] quit
[S2] interface gigabitethernet 1/0/4
[S2-GigabitEthernet1/0/4] port link-type trunk
[S2-GigabitEthernet1/0/4] port trunk allow-pass vlan 2 to 4
[S2-GigabitEthernet1/0/4] quit
S3 上的配置 ：部署 QinQ 功能，使上送到小区交换机的数据帧带有双层 VLAN标签。
<HUAWEI>system-view
[HUAWEI] sysname S3
[S3] vlan batch 201 401
[S3] interface gigabitethernet 1/0/1
[S3-GigabitEthernet1/0/1] port link-type trunk
[S3-GigabitEthernet1/0/1] port trunk allow-pass vlan 201 401
[S3-GigabitEthernet1/0/1] port vlan-stacking vlan 2 to 3 stack-vlan 201
[S3-GigabitEthernet1/0/1] port vlan-stacking vlan 4 stack-vlan 401
[S3-GigabitEthernet1/0/1] quit
[S3] interface gigabitethernet 1/0/2
[S3-GigabitEthernet1/0/2] port link-type trunk
[S3-GigabitEthernet1/0/2] port trunk allow-pass vlan 201 401
[S3-GigabitEthernet1/0/2] quit
S4 上的配置 ：部署 QinQ 功能，使上送到小区交换机的数据帧带有双层 VLAN标签。
<HUAWEI>system-view
[HUAWEI] sysname S4
[S4] vlan batch 201 401
[S4] interface gigabitethernet 1/0/1
[S4-GigabitEthernet1/0/1] port link-type trunk
[S4-GigabitEthernet1/0/1] port trunk allow-pass vlan 201 401
[S4-GigabitEthernet1/0/1] port vlan-stacking vlan 2 to 3 stack-vlan 201

```
[S4-GigabitEthernet1/0/1] port vlan-stacking vlan 4 stack-vlan 401
[S4-GigabitEthernet1/0/1] quit
[S4] interface gigabitethernet 1/0/2
[S4-GigabitEthernet1/0/2] port link-type trunk
[S4-GigabitEthernet1/0/2] port trunk allow-pass vlan 201 401
[S4-GigabitEthernet1/0/2] quit
```

S5上的配置：配置 2 to 1的VLAN映射功能。

```
<HUAWEI>system-view
[HUAWEI] sysname S5
[S5] vlan batch 501
[S5] interface gigabitethernet 1/0/1
[S5-GigabitEthernet1/0/1] port link-type trunk
[S5-GigabitEthernet1/0/1] port trunk allow-pass vlan 501
[S5-GigabitEthernet1/0/1] port vlan-mapping vlan 201 to 401 map-vlan 501
[S5-GigabitEthernet1/0/1] quit
[S5] interface gigabitethernet 1/0/2
[S5-GigabitEthernet1/0/2] port link-type trunk
[S5-GigabitEthernet1/0/2] port trunk allow-pass vlan 501
[S5-GigabitEthernet1/0/2] port vlan-mapping vlan 201 to 401 map-vlan 501
[S5-GigabitEthernet1/0/2] quit
[S5] interface gigabitethernet 1/0/3
[S5-GigabitEthernet1/0/3] port link-type trunk
[S5-GigabitEthernet1/0/3] port trunk allow-pass vlan 501
[S5-GigabitEthernet1/0/3] quit
```

完成上述配置后，不同家庭用户可以正常访问网络，且相同业务共用一个VLAN传输。

7.10 配置2 to 2的VLAN映射

2 to 2的VLAN映射功能可将数据帧中携带的双层VLAN标签映射为公网的双层VLAN标签。但配置 2 to 2的VLAN映射功能的端口类型也必须为Trunk或Hybrid，且端口必须加入映射后的公网VLAN中。

7.10.1 配置基于VLAN的2 to 2的VLAN映射

基于VLAN的 2 to 2的VLAN映射功能可实现端口在接收到带有VLAN标签的帧后，依据帧中的双层VLAN ID进行映射操作，将帧中的双层VLAN ID映射为指定的公网双层VLAN ID。

基于VLAN的 2 to 2的VLAN映射的具体配置方法很简单，具体如表 7-22所示（与7.9.1节介绍的 2 to 1的VLAN映射配置步骤差不多）。

表7-22 基于VLAN的2 to 2的VLAN映射的配置步骤

步骤	命令	说明
1	**system-view** 例如：<HUAWEI> **system-view**	进入系统视图
2	**interface** *interface-type interface-number* 例如：[HUAWEI]**interface gigabitethernet 0/0/1**	键入要配置 2 to 2 VLAN 映射的交换机端口，进入接口视图
3	**qinq vlan-translation enable** 例如：[HUAWEI-GigabitEthernet0/0/1] **qinq vlan-translation enable**	使能端口的 VLAN 转换功能。只有在端口使能了 VLAN 转换功能后，才可以在端口上配置 VLAN 映射和灵活 QinQ 功能。**本命令仅 S2700、S3700、S5700 和 S6700 系列需要配置** 缺省情况下，没有使能端口的 VLAN 转换功能，可用 **undo qinq vlan-translation enable** 命令取消端口的 VLAN 转换功能
4	**port vlan-mapping vlan** *vlan-id1* **inner-vlan** *vlan-id2* **map-vlan** *vlan-id3* **map-inner-vlan** *vlan-id4* [**remark-8021p** *8021p-value*] 例如：[HUAWEI-GigabitEthernet0/0/1] **port vlan-mapping vlan** 10 **inner-vlan** 20 **map-vlan** 100 **map-inner-vlan** 50	同时替换帧中的外层和内层 VLAN 标签。执行本命令可实现端口在接收到带有 VLAN 标签的帧后，依据帧中的 VLAN ID 进行映射操作，将帧中的双层 VLAN ID 映射为指定的双层公网 VLAN ID。命令中的各参数说明如下。 （1）*vlan-id1*：指定要映射的源外层标签的 VLAN ID，取值范围是 1～4 094 的整数 （2）*vlan-id2*：指定要映射的源内层标签的 VLAN ID，取值范围均为 1～4 094 的整数 （3）*vlan-id3*：指定映射后的外层标签的 VLAN ID，取值范围是 1～4 094 的整数

（续表）

步骤	命令	说明
4	**port vlan-mapping vlan** *vlan-id1* **inner-vlan** *vlan-id2* **map-vlan** *vlan-id3* **map-inner-vlan** *vlan-id4* [**remark-8021p** *8021p-value*] 例如：[HUAWEI-GigabitEthernet0/0/1] **port vlan-mapping vlan** 10 **inner-vlan** 20 **map-vlan** 100 **map-inner-vlan** 50	（4）*vlan-id4*：指定映射后的内层标签的 VLAN ID，取值范围均为 1～4 094 的整数 （5）*8021p-value*：可选参数，指定修改映射后的 VALN 标签的 802.1p 优先级，取值范围为 0～7，值越大优先级越高。选择此可选参数可实现端口在接收到带标签的帧后，将帧中的 802.1p 优先级修改为用户配置的 802.1p 优先级值 缺省情况下，交换机端口下没有配置对数据帧中携带的 VLAN 标签进行映射操作，可用 **undo port vlan-mapping vlan** *vlan-id1* **inner-vlan** *vlan-id2* **map-vlan** *vlan-id3* **map-inner-vlan** *vlan-id4* [**remark-8021p** *8021p-value*]取消替换带有指定双层 Tag 的帧的双层标签 VLAN 映射操作

【示例 】配置2 to 2的VLAN映射功能，将数据帧中携带的两层标签（外层标签为10、内层标签为20）映射为外层标签为100、内层标签为200，再进行转发。

```
<HUAWEI>system-view
[HUAWEI] interface gigabitethernet1/0/1
[HUAWEI-GigabitEthernet1/0/1] port link-type trunk
[HUAWEI-GigabitEthernet1/0/1] port trunk allow-pass vlan 100
[HUAWEI-GigabitEthernet1/0/1] port vlan-mapping vlan 10 inner-vlan 20 map-vlan 100 map-inner-vlan 200
```

7.10.2 配置基于流策略的2 to 2的VLAN映射

整体来说，基于流策略的 2 to 2的VLAN映射的配置方法与 7.9.2节介绍的基于流策略的 2 to 1的VLAN映射的配置方法差不多，只是多了一个内层标签的重标记动作。具体如表7-23所示（可在端口的入方向或者出方向上应用 ）。

表7-23 基于流策略的2 to 2的VLAN映射的配置步骤

配置任务	步骤	命令	说明
定义流分类	1	**system-view** 例如：< HUAWEI > **system-view**	进入系统视图
	2	**traffic classifier** *classifier-name* **operator and** 例如：[HUAWEI]**traffic classifier** c1	创建流分类并进入类视图，并指定流分类下各规则之间是逻辑"与"的关系。其他说明参见表 7-21 的第 2 步
	3	**if-match vlan-id** *vlan-id* 例如：[HUAWEI-classifier-c1]**if-match vlan-id** 2	配置匹配数据帧的规则，即指定数据帧的源外层 VLAN ID。其他说明参见表 7-21 中的第 3 步
	4	**if-match cvlan-id** *cvlan-id* 例如：[HUAWEI-classifier-c1] **if-match cvlan-id** 20	配置匹配数据帧的规则，即指定数据帧的源内层 VLAN ID。参数 *cvlan-id* 用来指定匹配的内层 VLAN ID，取值范围均为 1～4 094 的整数 缺省情况下，流分类中没有基于 QinQ 数据帧内外两层 VLAN ID 进行分类的匹配规则，可用 **undo if-match cvlan-id** *vlan-id2* 命令在流分类中删除指定的匹配规则
	5	**quit** 例如：[HUAWEI-classifier-c1]**quit**	退出流分类视图，返回系统视图

（续表）

配置任务	步骤	命令	说明
定义流行为	6	**traffic behavior** *behavior-name* 例如：[HUAWEI] **traffic behavior** b1	创建流行为并进入流行为视图。其他说明参见表 7-15 的第 5 步
	7	**remark vlan-id** *vlan-id* 例如：[HUAWEI-behavior-b1]**remark vlan-id** 200	配置流行为，在流行为中创建重标记帧中的外层 VLAN ID 标签。其他说明参见表 7-19 中的第 6 步
	8	**remark cvlan-id** *vlan-id* 例如：[HUAWEI-behavior-b1] **remark cvlan-id** 100	配置流行为，在流行为中创建重标记帧中的内层 VLAN ID 标签。参数 *vlan-id* 用来指定重标记数据帧的内层 VLAN ID 值，取值范围为 1～4 094 的整数 缺省情况下，流行为中没有重标记 VLAN 数据帧的内层 VLAN 标签值的动作，可用 **undo remark cvlan-id** 命令恢复缺省情况
	9	**quit** 例如：[HUAWEI-behavior-b1] **quit**	退出流行为视图，返回系统视图
创建 QoS 策略，关联流分类与流行为	10	**traffic policy** *policy-name* 例如：[HUAWEI] **traffic policy** p1	创建流策略并进入流策略视图。其他说明参见表 7-15 的第 8 步
	11	**classifier** *classifier-name* **behavior** *behavior-name* 例如：[HUAWEI-trafficpolicy-p1]**classifier** c1 **behavior** b1	将以上定义的流分类与指定的流行为进行绑定，组成流策略。其他说明参见表 7-15 的第 9 步
	12	**quit** 例如：[HUAWEI-dsdomain-ds1] **quit**	退出流策略视图，返回系统视图
在交换机端口入或出方向上应用 QoS 策略	13	**interface** *interface-type interface-number* [HUAWEI] **interface gigabitethernet 1/0/1**	键入要应用流策略的交换机端口，进入接口视图
	14	**port link-type hybrid** 例如：[HUAWEI-GigabitEthernet1/0/1] **port link-type hybrid**	配置以上端口的类型为 Hybrid
	15	**port hybrid Untagged vlan** { { *vlan-id1* [**to** *vlan-id2*] } &<1-10> | **all** } 例如：[HUAWEI-GigabitEthernet1/0/1] **port hybrid Untagged vlan** 2 3	把以上 Hybrid 端口以不带标签方式加入到映射后的外层 VLAN 中，具体命令参数本第 6 章已有详细介绍，不再赘述
	16	**traffic-policy** *policy-name* { **inbound** | **outbound** } 例如：[HUAWEI-GigabitEthernet1/0/1] **traffic-policy p1 inbound**	在端口的入或出方向应用流策略。其他说明参见表 7-19 的第 14 步

7.10.3 基于VLAN的2 to 2的VLAN映射配置示例

本示例拓扑结构如图7-25所示。处于不同地理位置的用户为了便于用户自己规划的私网VLAN ID不与ISP网络中的VLAN ID冲突，采用了QinQ方式传输，使用户数据帧中带有双层VLAN标签。但这同时也带来了一个因用户数据帧中的VLAN ID与ISP网络分配的VLAN ID不一致导致用户数据帧被丢弃的问题，从而导致CE两端的用户无法正常通信。此时可采用 2 to 2的VLAN映射解决这一问题，实现 CE Switch5 和Switch6中的用户互通。

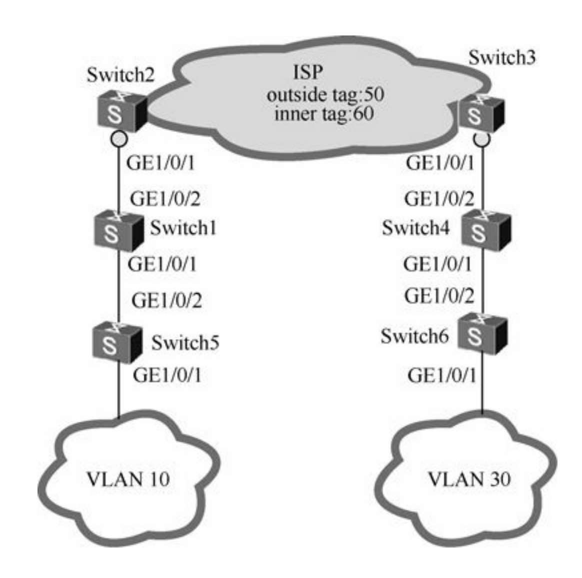


图7-25 基于VLAN的2 to 2的VLAN映射配置示例的拓扑结构

本示例采用的是基于VLAN的2 to 2的VLAN映射解决方案。

1．配置思路分析

本示例的基本配置思路如下。

（1）在连接用户的Switch5和Switch6交换机上将端口加入VLAN 10或VLAN 30中。

（2）在Switch1和Switch4上配置QinQ，使发往ISP网络的数据帧带有双层VLAN标签。

（3）在连接 ISP网络的Switch2和Switch3交换机上部署 2 to 2的VLAN映射功能。

2．具体配置步骤

下面是本示例的具体配置步骤。

（1）将连接用户的交换机的下行口划分到指定的VLAN中。

Switch5上的配置：

```
<HUAWEI>system-view
[HUAWEI] sysname Switch5
[Switch5] vlan 10
[Switch5-vlan10] quit
[Switch5] interfacegigabitethernet 1/0/1
[Switch5-GigabitEthernet1/0/1] port link-type access
[Switch5-GigabitEthernet1/0/1] port default vlan 10
[Switch5] interface gigabitethernet 1/0/2
[Switch5-GigabitEthernet1/0/2] port link-type trunk
```

[Switch5-GigabitEthernet1/0/2] port trunk allow-pass vlan 10

Switch6上的配置：

<HUAWEI>system-view

[HUAWEI] sysname Switch6

[Switch6] vlan 30

[Switch6-vlan30] quit

[Switch6] interfacegigabitethernet 1/0/1

[Switch6-GigabitEthernet1/0/1] port link-type access

[Switch6-GigabitEthernet1/0/1] port default vlan 30

[Switch6] interfacegigabitethernet 1/0/2

[Switch6-GigabitEthernet1/0/2] port link-type trunk

[Switch6-GigabitEthernet1/0/2] port trunk allow-pass vlan 30

（2）在Switch1和Switch4上配置QinQ，使发往ISP网络的数据帧带有双层VLAN标签。

Swtich1上的配置：

<HUAWEI>system-view

[HUAWEI] sysname Swtich1

[Swtich1] vlan 20

[Swtich1-vlan20] quit

[Switch1] interfacegigabitethernet 1/0/1

[Switch1-GigabitEthernet1/0/1] port link-type trunk

[Swtich1-GigabitEthernet1/0/1] port trunk allow-pass vlan 20

[Swtich1-GigabitEthernet1/0/1] port vlan-stacking vlan 10 stack-vlan 20

[Swtich1-GigabitEthernet1/0/1] quit

[Swtich1] interfacegigabitethernet 1/0/2

[Switch1-GigabitEthernet1/0/2] port link-type trunk

[Swtich1-GigabitEthernet1/0/2] port trunk allow-pass vlan 20

[Swtich1-GigabitEthernet1/0/2] quit

Swtich4上的配置：

<HUAWEI>system-view

[HUAWEI] sysname Swtich4

[Swtich4] vlan 40

[Swtich4-vlan40] quit

[Swtich4] interfacegigabitethernet 1/0/1

[Switch4-GigabitEthernet1/0/1] port link-type trunk

[Swtich4-GigabitEthernet1/0/1] port trunk allow-pass vlan 40

[Swtich4-GigabitEthernet1/0/1] port vlan-stacking vlan 30 stack-vlan 40

[Swtich4-GigabitEthernet1/0/1] quit

[Swtich4] interface gigabitethernet 1/0/2

[Switch4-GigabitEthernet1/0/2] port link-type trunk

[Swtich4-GigabitEthernet1/0/2] port trunk allow-pass vlan 40

[Swtich4-GigabitEthernet1/0/2] quit

（3）在连接 ISP网络的Switch2和Switch3上部署 2 to 2 的VLAN映射功能。

Swtich2上的配置：

<HUAWEI>system-view

[HUAWEI] sysname Swtich2

[Swtich2] interface gigabitethernet 1/0/1

[Swtich2-GigabitEthernet1/0/1] port vlan-mapping vlan 20 inner-vlan 10 map-vlan 50 map-inner-vlan 60

Swtich3上的配置 ：

<HUAWEI>system-view

[HUAWEI] sysname Swtich3

[Swtich3] interfacegigabitethernet 1/0/1

[Swtich3-GigabitEthernet1/0/1] port vlan-mapping vlan 40 inner-vlan 30 map-vlan 50 map-inner-vlan 60

完成上述配置后，CE1中的用户就可以与CE2中的用户互通了，因为此时他们发出的数据帧中的双层VLAN标签经过Switch2或Switch3后是一致的了，即内层标签为60，外层标签为50。

7.10.4 基于流策略的2 to 2的VLAN映射配置示例

本示例拓扑结构如图 7-26所示，企业A和企业B各自规划自己的私网VLAN ID，但是由于用户数据帧中的VLAN ID与 ISP网络分配的VLAN ID不一致，将导致用户数 据帧被丢弃，从而导致用户通信中断。此时可在CE侧交换机上部署VLAN映射功能，实现企业A与企业B通过运营商网络互相通信。本示例采用的是基于流策略的2 to 2的VLAN映射解决方案。

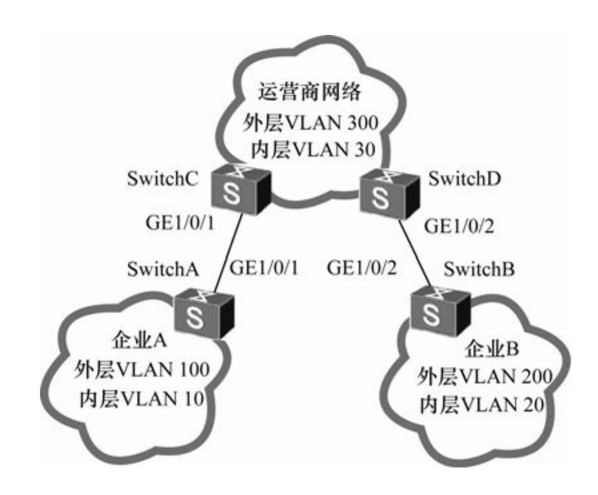


图7-26 基于流策略的2 to 2的VLAN映射的配置示例的拓扑结构

1．配置思路分析

本示例的基本配置思路如下。

（1）在SwitchA、SwitchB、SwitchC、SwitchD上创建各自所属的外层VLAN。

（2）在 SwitchA 和 SwitchB 上创建各自的类、流行为、流策略。

（3）配置SwitchA、SwitchB、SwitchC、SwitchD接口加入各自创建的VLAN。

（4）配置SwitchA上GE1/0/1端口和SwitchB上GE1/0/2端口应用基于流策略的替换双层标签的VLAN映射功能。

2．具体配置步骤

（1）在各交换机上创建所需的外层VLAN，并将对应的端口加入这些VLAN中。

SwitchA上的配置：

<HUAWEI>system-view

[HUAWEI] sysname SwitchA

[SwitchA] vlan 100

[SwitchA] quit

[SwitchA] interfacegigabitethernet 1/0/1

[SwitchA-GigabitEthernet1/0/1] port link-type trunk

[SwitchA-GigabitEthernet1/0/1] port trunk allow-pass vlan 100

SwitchB上的配置 ：

<HUAWEI>system-view

[HUAWEI] sysname SwitchB

[SwitchB] vlan 200

[SwitchB] quit

[SwitchB] interface gigabitethernet 1/0/2

[SwitchB-GigabitEthernet1/0/2] port link-type trunk

[SwitchB-GigabitEthernet1/0/2] port trunk allow-pass vlan 200

SwitchC上的配置 ：

<HUAWEI>system-view

[HUAWEI] sysname SwitchC

[SwitchC] vlan 300

[SwitchC] quit

[SwitchC] interfacegigabitethernet 1/0/1

[SwitchC-GigabitEthernet1/0/1] port link-type trunk

[SwitchC-GigabitEthernet1/0/1] port trunk allow-pass vlan 300

SwitchD上的配置：

<HUAWEI>system-view

[HUAWEI] sysname SwitchD

[SwitchD] vlan 300

[SwitchD] quit

[SwitchD] interfacegigabitethernet 1/0/2

[SwitchD-GigabitEthernet1/0/2] port link-type trunk

[SwitchD-GigabitEthernet1/0/2] port trunk allow-pass vlan 300

（2）在SwitchA和SwitchB的入、出方向上同时配置流分类、流行为、流策略。

SwitchA入方向的配置：

[SwitchA] traffic classifier name1 operator and

[SwitchA-classifier-name1] if-match vlan-id 300

[SwitchA-classifier-name1] if-match cvlan-id 30

[SwitchA-classifier-name1] quit

[SwitchA] traffic behavior name1
[SwitchA-behavior-name1] remark vlan-id 100
[SwitchA-behavior-name1] remark cvlan-id 10
[SwitchA-behavior-name1] quit
[SwitchA] traffic policy name1
[SwitchA-trafficpolicy-name1] classifiername1 behaviorname1
SwitchA出方向的配置：
[SwitchA] traffic classifier name2 operator and
[SwitchA-classifier-name2] if-match vlan-id 100
[SwitchA-classifier-name2] if-match cvlan-id 10
[SwitchA-classifier-name2] quit
[SwitchA] traffic behavior name2
[SwitchA-behavior-name2] remark vlan-id 300
[SwitchA-behavior-name2] remark cvlan-id 30
[SwitchA-behavior-name2] quit
[SwitchA] traffic policy name2
[SwitchA-trafficpolicy-name2] classifiername2 behaviorname2
SwitchB入方向的配置：
[SwitchB] traffic classifier name1 operator and
[SwitchB-classifier-name1] if-match vlan-id 300
[SwitchB-classifier-name1] if-match cvlan-id 30
[SwitchB-classifier-name1] quit
[SwitchB] traffic behaviorname1
[SwitchB-behavior-name1] remark vlan-id 200
[SwitchB-behavior-name1] remark cvlan-id 20
[SwitchB-behavior-name1] quit
[SwitchB] traffic policy name1
[SwitchB-trafficpolicy-name1] classifiername1 behavior name1
SwitchB出方向的配置：
[SwitchB] traffic classifier name2 operator and
[SwitchB-classifier-name2] if-match vlan-id 200
[SwitchB-classifier-name2] if-match cvlan-id 20
[SwitchB-classifier-name2] quit
[SwitchB] traffic behaviorname2
[SwitchB-behavior-name2] remark vlan-id 300
[SwitchB-behavior-name2] remark cvlan-id 30
[SwitchB-behavior-name2] quit
[SwitchB] traffic policy name2
[SwitchB-trafficpolicy-name2] classifiername2 behaviorname2
（3）在SwitchA和SwitchB上配置基于流策略的替换双层标签的VLAN映射功能。

SwitchA上的配置：

<SwitchA>system-view

[SwitchA] interfacegigabitEthernet 1/0/1

[SwitchA-GigabitEthernet1/0/1] traffic-policy name1 inbound

[SwitchA-GigabitEthernet1/0/1] traffic-policy name2 outbound

SwitchB上的配置：

<SwitchB>system-view

[SwitchB] interfacegigabitEthernet 1/0/2

[SwitchB-GigabitEthernet1/0/2] traffic-policy name1 inbound

[SwitchB-GigabitEthernet1/0/2] traffic-policy name2 outbound

完成以上配置后，企业A内用户与企业B内用户就可以互相访问了。

第8章　生成树协议配置与管理

生成树协议就是用来消除网络中可能存在的二层环路，以防广播风暴，或者数据传输死循环。在我们日常进行的小型局域网交换机配置中不是很常用，因为一般很少出现二层环路。但是它在一些结构比较复杂，网络规模比较大的网络中确实是必需之选，如为了解决单台交换机的单点故障问题而特意添加的冗余链路（特别是在汇聚层和核心层中），还有为了实现VLAN流量负载分担而特意部署的多条等价链路等，以及为了提高网络的可靠性，在一些网络中存在接入环网等。

华为S系列交换机全面支持STP（生成树协议）、RSTP（快速生成树协议）和MSTP（多生成树协议）这三种生成树协议，全面满足于各类网络环境下的生成树计算。STP是最原始的生成树协议，它的主要不足是网络拓扑收敛速度慢，RSTP是在STP基础上改进的生成树协议版本，它的收敛速度明显提高。但无论是STP还是RSTP都是单生成树，即一个网络中只生成一棵生成树。MSTP又是在RSTP基础上的改进版本，最主要的特性就是可以在一个网络中划分多个MST域，在一个MST域中又可划分多个生成树实例，是多生成树协议。

本章将全面介绍华为S系列交换机的STP、RSTP和MSTP功能的配置。总体上配置都很简单，仅包括基本功能和影响拓扑收敛性能的参数配置，在RSTP和MSTP中还可配置一些保护功能。

8.1 STP基础

STP（Spanning Tree Protocol，生成树协议）是根据 IEEE 802.1D标准建立的，用于在局域网中消除数据链路层环路的协议。运行STP协议的设备通过彼此交互信息发现网络中的环路，并有选择地对某些端口进行阻塞，以最终实现将环路网络结构修剪成无环路的树型网络结构，从而防止报文在环路网络中不断增生和无限循环，避免设备由于重复接收相同的报文所造成的报文处理能力下降的问题发生。

8.1.1 STP的由来

说到STP（包括后面将要介绍的RSTP和MSTP），许多读者朋友一直想不明白它有什么用，因为在常见的星型以太网中，通常很少需要配置它，网络照样可以正常使用。的确如此，因为在由星型结构单元组成的以太网中通常就已是无交叉、无物理环路的树形结构。那么STP到底有什么用，主要在什么情况下使用它呢？在此我要告诉你，STP仅在网络中存在冗余链路，或者网络中存在环形网络结构（其实都是存在封闭的物理环路）时才需要采用，其目的就是消除网络这些可能造成数据往返传输，形成死循环的冗余链路，以及消除同一个交换机从不同端口上收到多份相同的数据。

在以太网交换网络中为了进行链路备份，提高网络可靠性，通常会在一些关键设备间使用冗余链路，如在图8-1所示两核心交换机使用了冗余连接。但是使用冗余链路会在交换网络上产生环路（如图中的S1和

S2的port1、port2端口就构成了一个物理封闭环路），并导致广播风暴以及MAC地址表不稳定等故障现象，从而导致用户通信质量较差，甚至通信中断。如S1的port1端口在从S2的port1端口收到一个广播包后，会再从包括它的port2在内的所有其他端口发送出去，这样S2的port2端口又会收到这个同样的广播包，然后从包括它的port1在内的所有其他端口发送出去，致使S1的port1端口再次收到同样的广播包，就这样一直循环下去，形成了一个死循环，最终形成广播风暴。

另外，还有一些网络中存在设备间的封闭环形结构，如在图8-2所示的网络中，S1、S2、S3和LAN局域网所形成的封闭环。其实这种结构和图8-1是一样的，只是相当于在图8-1中的直连链路中间加了一些设备S2和S3，同样会形成广播风暴。使用STP技术，其实更多是抱着以防万一的心态，怕网络中存在这样的物理封闭环路。因为STP技术在保障正常使用冗余链路备份的同时，又可确保不会出现二层通信环路。

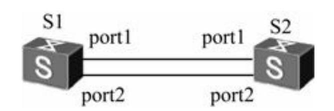


图8-1 冗余链路结构示例

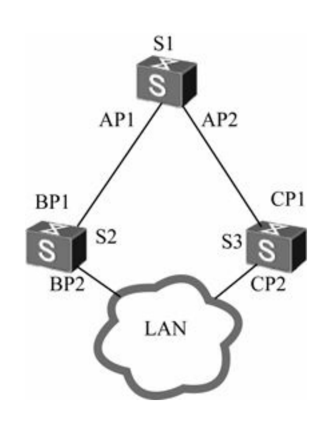


图8-2 环形网络结构示例

为解决交换网络中的环路问题，提出了生成树协议STP。运行STP协议的设备通过彼此交互信息发现网络中的环路，并有选择地对某个端口进行阻塞，最终将环形网络结构修剪成无环路的树形网络结构，从而防止报文在环形网络中不断循环，避免设备由于重复接收相同的报文造成处理能力下降。

再如图8-3所示环形结构网络中，如果没有在交换机上启用STP协议就会产生如下两种情况。

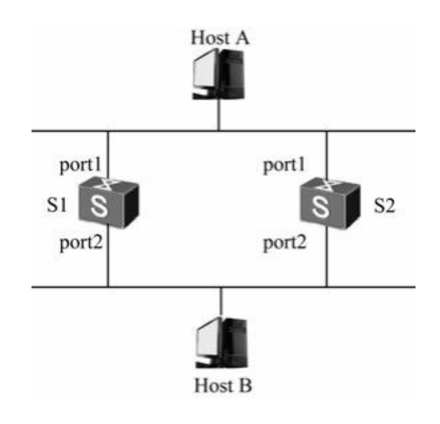


图8-3 环形结构网络示例

（1）广播风暴导致网络不可用

如果HostA发出广播请求，交换机S1和S2的端口port1都将收到这个广播报文，然后从所有其他同网段的端口（如 port2）广播出去，这时对端交换机与之相连的端口又收到相同的广播报文，这台交换机再从它的其他端口（如 port1）转发出去，对端交换机又会收到相同的广播报文，如此反复，最终导致整个网络资源被耗尽，网络瘫痪不可用。

（2）MAC地址表震荡导致MAC地址表项被破坏

即使是单播报文，也有可能导致交换设备的MAC地址表项混乱，以致破坏交换设备的MAC地址表。假设图8-3所示的网络中没有广播风暴，HostA发送一个单播报文给HostB，如果此时HostB临时从网络中移去，那么交换机上有关HostB的MAC地址表项也将被删除。此时HostA发给HostB的单播报文，将被交换机S1和S2上的端口port1接收，由于S1上没有相应的MAC地址转发表项了，因此该单播报文将被同时泛洪转发到其他端口（如port2）上，交换机S2的端口port2在收到从对端port2端口发来的单播报文后，然后又以泛洪方式从其他端口（如 port1）发出去，使交换机 S1 的 port1 端口又会收到这个单播报文。如此反复，在两台交换机上，由于不间断地从端口 port2、port1 收到主机 A 发来的单播报文，交换机会不停地修改自己的MAC地址表项，从而引起了MAC地址表的抖动。如此下去，最终导致MAC地址表项被破坏。

8.1.2 STP基本概念

在STP协议中涉及许多基本概念，如根桥、桥 ID、桥优先级、根端口、指定端口、端口状态、端口ID、端口优先级等。所以这些基本概念可以用"一个根桥、两种度量、三个选举要素、四个比较原则和五种端口状态"一句话来形容，下面分别予以介绍。

1．一个根桥

树形的网络结构必须有一个树根，于是STP引入了根桥（Root Bridge）概念。对于一个运行STP协议的网络，根桥在全网中只有一个，就像一棵树只有一个树根一样，那就是网络中具有最小桥ID（BID）的桥。网络中除根桥外的其他桥统称为非根桥。有关桥ID将在下面具体介绍。

说明

根桥是整个网络的逻辑中心，但不一定是物理中心，且会根据网络拓扑的变化而 动态变化。一般是需要将环路中所有交换机当中性能最好的一台设置为根桥交换机，以保证能够提供最好的网络性能和可靠性。网络收敛后，根桥会按照一定的时间间隔产生并向外发送配置 BPDU，其他设备仅对该报文进行转发，传达拓扑变化记录，从而保证拓扑的稳定。

2．两种度量

在 STP 生成树的计算中要确定两个方面：一是确定哪台交换机将成为根桥，在非根桥的交换机中哪些端口具有收、发数据的功能，哪些端口又该被阻塞，以便最终形成无环路的树形结构交换网络。这里的STP生成树计算所依据的就是STP中的ID和路径开销。

（1）ID。STP中的 ID包括：BID（Bridge ID，桥 ID）和PID（Port ID，端口 ID）两种。

① BID。BID一共 64位，高 16位为桥优先级（Bridge Priority）值，低 48位为桥背板MAC地址。BID决定了哪台交换机将成为交换网络中的根桥。在STP中规定BID最小的交换机将被选举为根桥。在进行根桥选举中，首先要比较的就是高16位的桥优先级，它是一个用户可以设定的参数，数值范围从0到61440，设定的值越小，优先级越高，也越有可能成为根桥。如果各交换机的桥优先级都一样才比较它们BID中的桥背板MAC地址了，MAC地址最小的将成为该交换网络中的根桥。

② PID。PID由两部分构成，高 4位是端口优先级，低 12位是端口号。PID只在某些情况下对选择"指定端口"有作用。即在选择指定端口时，两个端口的根路径开销和发送BPDU交换机的BID都相同的情况下，

比较端口的PID，PID小者为指定端口。端口优先级可以影响端口在指定生成树实例上的角色。

（2）路径开销。路径开销（Path Cost）是一个端口参数，由具体端口的链路速率决定（对于聚合链路，链路速率是聚合组中所有状态为Up的成员口的速率之和），是STP协议用于选择链路的参考值。STP协议通过计算各端口的路径开销，选择较为“强壮”的链路，阻塞多余的链路，将网络修剪成无环路的树形网络结构。

在一个运行STP协议的交换网络中，某端口到根桥累计的路径开销就是所经过的各个桥上的各端口的路径开销累加值，这个值叫做根路径开销（Root Path Cost）。根桥上所有端口的根路径开销，以及同交换机上不同端口间的路径开销值均为零。

3．三个选举要素

由环形网络拓扑结构修剪为树形结构，需要使用STP中的3个选举要素，即根桥、根端口和指定端口。下面结合图8-4进行介绍。

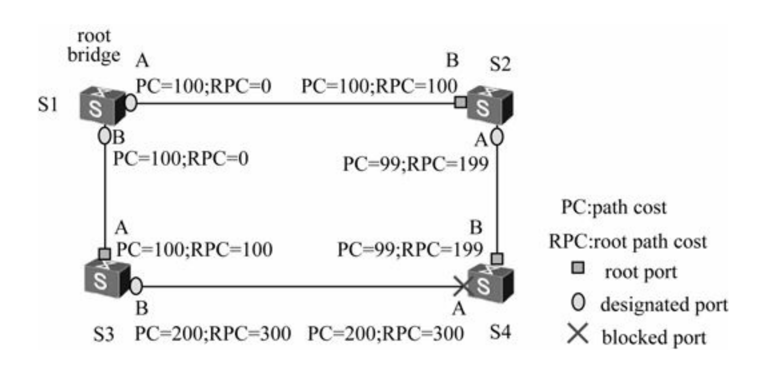


图8-4 STP三要素描述示例

图中的PC为Past Cost，即“路径开销”，RPC为Root Path Cost，即“根路径开销”。

（1）根桥

根桥（Root Bridge，RB）就是BID最小的桥设备，通过交互配置BPDU报文来选出最小的BID，如图8-4中的S1为根桥。这部分已在前面有了详细介绍，不再赘述。

（2）根端口

根端口（Root Port，RP）负责向根桥方向转发数据，是当前桥设备上去往根桥的“根路径开销”（Root Path Cost，RPC）值最小的端口，也即非根桥的交换机上离根桥“最近”的端口。

当多个端口根路径开销相同时，会先比较指定桥ID，对应最小指定桥ID的端口会成为根端口；当指定桥ID也相同时，才会比较指定桥上的PID，指定桥上具有最小PID的对应端口会成为根端口。端口优先级的取值范围为0到255，值越小，端口的优先级就越高。很显然，在一个运行**STP**协议的设备上根端口有且只有一个，但根桥上没有根端口 。

通过比较图8-4中S2、S3和S4中的A、B端口到达根桥S1的根路径开销（RPC）值就可以得出它们的根端口了（分别为B端口、A端口和B端口）。

（3）指定端口

“指定端口”（Designated Port，DP）与“指定桥”（Designated Bridge，DB）息息相 关，但不是一一对应的，如何确定要分以下两种情况。

① 对于一台设备而言，与本机直接相连并且负责向本机转发配置消息的设备就是指定桥，指定桥中向本桥转发配置消息的端口就是指定端口。

② 对于一个局域网而言，负责向本网段转发配置消息的设备就是指定桥，指定桥上向本网段转发配置

消息的端口就是指定端口。

如图8-5所示，AP1、AP2、BP1、BP2、CP1、CP2分别表示设备S1、S2、S3的端口。S1通过端口AP1向S2转发配置消息，则S2的指定桥就是S1，指定端口就是S1的端口AP1。而与局域网LAN相连的有S2和S3 两台设备，如果配置S2负责向LAN转发配置消息，则LAN的指定桥就是S2，指定端口就是S2的BP2。

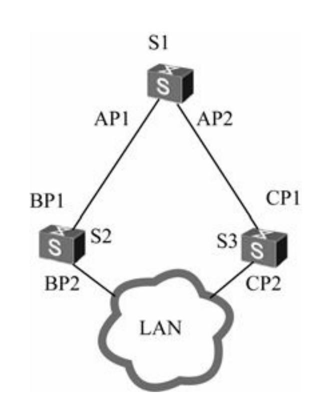


图8-5 指定桥与指定端口示例

一旦根桥、根端口、指定端口选举成功，则整个树形拓扑建立完毕。在拓扑稳定后，只有根端口和指定端口转发流量，其他的非根、非指定端口（称为阻塞端口）都处于阻塞（Blocking）状态，它们只接收STP协议报文而不转发用户流量，如图8-4中的S4的A端口。

4．四个比较原则

STP 生成树的生成计算中依据的就是各个端口在发送配置 BPDU 中所携带的 4个优先级向量：{ 根桥ID，累计根路径开销，发送者BID，发送端口 PID }。具体解释如下。

（1）根桥ID：每个STP网络中有且仅有一个根。

（2）累计根路径开销：发送配置BPDU的端口到根桥的距离。

（3）发送者BID：发送配置BPDU的设备的BID。

（4）发送端口PID：发出配置BPDU的端口的PID。

STP网络中的其他设备收到配置BPDU消息后，将比较这些字段值，然后按照以下4个基本比较原则（在STP计算过程中，遵循数值越小越好的原则）。

（1）最小BID：用来选举根桥。运行STP协议的设备之间根据各自发送的配置BPDU中BID字段值最小的作为根桥。根桥的选举原则是通过BID中的桥优先级和桥MAC地址进行比较的，先进行桥优先级比较，优先级最高（优先级值最小）的将成为根桥；桥优先级相同再比较桥MAC地址，MAC地址最小的将成为根桥。

（2）最小累计根路径开销：用来在非根桥上选择根端口。在运行STP协议的设备上到达根桥的总路径开销值最小的端口作为该桥的根端口。在根桥上，每个端口到根桥的根路径开销都是0，所以根端口都是在指定桥上的，而不是在根桥上的 。

（3）最小发送者BID：用来在非根桥上选择指定桥和根端口。当一台运行STP协议的设备要在两个以上根路径开销相等的非根桥中选择指定桥，要在接收配置BPDU的多个端口之中选择根端口时，通过 STP协议计算将选择接收到的配置 BPDU中发送者的 BID 较小的那个桥作为自己的指定桥，接收该配置 BPDU的端口就作为自己的根端口。

如图8-4所示，假设S2的BID小于S3的BID，如果S4的A、B两个端口接收到的BPDU里面的根路径开销相等，那么S2将成为 S4的指定桥，S4端口B将成为S4的根端口。

（4）最小PID：用于在根路径开销相同的情况下，阻塞PID值较大的端口，PID值最小的端口将成为该桥上的指定端口。在如图8-6所示的情况下PID才起作用，S1的端口A的PID小于端口B的PID，由于两个端口上收到的BPDU中根路径开销、发送设备的BID都相同，所以消除环路的依据就只有PID。

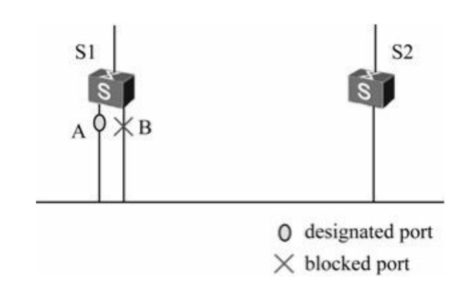


图8-6 根据PID选择指定端口的示例

5．五种端口状态

运行STP协议的设备上有以下5种端口状态。

（1）Forwarding：转发状态，此时端口既转发用户流量也转发BPDU报文。只有根端口或指定端口才能进入Forwarding状态。

（2）Learning：学习状态，此时设备会根据收到的用户流量构建MAC地址表，但不转发用户流量。这是一种过渡状态，增加Learning状态为防止临时二层环路。

（3）Listening：监听状态，此时设备正在确定端口角色，将选举出根桥、根端口和指定端口。这也是一种过渡状态。

（4）Blocking：阻塞状态，此时端口仅可接收并处理BPDU，不转发用户流量。

（5）Disabled：禁用状态，此时端口不仅不能转发 BPDU 报文，也不能转发用户流量。端口状态为Down。

以上这5种端口状态的迁移机制如图8-7所示。图中的序号说明如下。

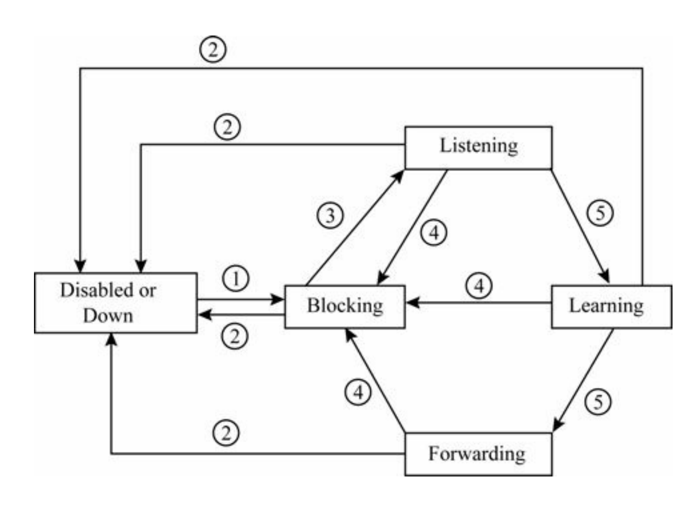


图8-7 端口状态迁移示意图

①：端口从禁止状态开始初始化或者使能后首先进入阻塞状态。

②：在端口突然被禁用或者链路失效时将从当前其他所有状态下直接进入到禁用状态。

③：在端口被选举为根端口或指定端口后，由阻塞状态进入到监听状态。

④：在端口不再是根端口或指定端口时，会从当前其他状态直接进入阻塞状态。

⑤：当新选出的根端口和指定端口要经过两倍的转发延时（即从监听状态进入学习状态，再从学习状态进入转发状态）后才能进入转发状态，以确保新的配置消息传遍整个网络，防止临时环路的产生。

说明

华为公司数据通信设备缺省情况处于 MSTP 模式，当从 MSTP 模式切换到 STP 模式，运行STP协议的设备上端口支持的端口状态仍然保持和MSTP支持的端口状态一样（MSTP端口状态与RSTP端口状态相同），支持的状态仅包括Forwarding（转发）、Learning（学习）和Discarding（丢弃）这三种，将在本章后面具体介绍。

8.1.3 STP的3个定时器

对于STP，影响端口状态和端口收敛有以下3个定时器参数。

1．Hello Time（Hello定时器）

Hello定时器是指运行STP协议的设备发送配置BPDU的时间间隔，即设备会每隔Hello Time时间向周围的设备发送配置消息BPDU，以确认链路是否存在故障。

当网络拓扑稳定之后，该定时器的修改只有在根桥修改后才有效。新的根桥会在发出的BPDU报文中填充适当的字段以向其他非根桥传递该定时器修改的信息。当拓扑变化之后，TCN BPDU的发送不受这个定时器的管理。

2．Forward Delay（转发延时）

转发延时是设备进行状态迁移的延迟时间，是指一个端口处于Listening和Learning状态的各自持续时间，缺省是15s。即Listening状态持续15s，随后进入Learning状态，然后再持续15s。

链路故障会引发网络重新进行生成树的计算，生成树的结构将发生相应的变化。 但是重新计算得到的新配置消息不可能立即传遍整个网络，如果此时新选出的根端口和指定端口就立即开始数据转发的话很可能会造成临时的二层环路。为此，STP 采用了一种状态迁移机制，新选出的根端口和指定端口要经过两倍的 Forward Delay延时后才能进入转发状态，这个延时保证了新的配置消息传遍整个网络，从而防止了临时环路的产生。

3．Max Age（最大生存时间）

最大生存时间是指端口的BPDU报文的老化时间，可在根桥上通过命令人为改动老化时间。

Max Age通过配置BPDU报文的传输，可保证Max Age在整网中一致。运行 STP协议的网络中非根桥设备收到配置BPDU报文后，会对报文中的Message Age（消息生存时间）和Max Age进行比较：如果Message Age小于等于Max Age，则该非根桥设备继续转发配置BPDU报文；如果Message Age大于Max Age，则该配置BPDU报文将被老化。该非根桥设备直接丢弃该配置 BPDU，可认为网络直径过大，导致根桥连接失败。

说明

当配置 BPDU从根桥发出时报文中的Message Age值为 0。当其他桥收到配置BPDU时，Message Age值为从根桥发送到当前桥接收到BPDU的总时间，包括传输延时等。实际实现中，配置BPDU报文每经过一个桥，Message Age增加 1。

8.1.4 STP BPDU报文

STP协议采用的是BPDU（Bridge Protocol Data Unit，桥协议数据单元）类型报文，也称为配置消息。STP就是通过在设备之间传递BPDU来确定最终修剪完成的网络拓扑结构。STP BPDU又分为两大类。

（1）配置BPDU（Configuration BPDU）：用来进行生成树计算和维护生成树拓扑的报文，是初始阶段

中各交换机发送的BPDU消息。

（2）TCN BPDU（Topology Change Notification BPDU）：当拓扑结构发生变化时，下游设备用来通知上游设备网络拓扑结构发生变化的报文。它是当拓扑稳定后，网络中出现了链路故障，网络拓扑发生改变时所发送的BPDU消息。

“配置 BPDU”是一种心跳报文，只要端口使能 STP 协议，则设备就会按照Hello Time定时器规定的时间间隔从指定端口发送配置 BPDU；而 TCN BPDU是在设备检测到网络拓扑发生变化时才发出的。STP BPDU报文被封装在以太网数据帧中，此时目的MAC地址是组播MAC地址：01-80-C2-00-00-00，在LLC头部中IEEE为STP保留的DSAP和SSAP值均为0x42，Control为0x03。下面具体介绍这两种BPDU。

1．配置BPDU

在STP中通常所说的BPDU报文多数指配置BPDU。在初始化过程中，每个桥都主动发送配置BPDU。但在网络拓扑稳定以后，只有根桥主动发送配置**BPDU**，其他桥在收到上游传来的配置BPDU后才触发发送自己的配置BPDU。具体来说，配置BPDU在以下3种情况下会产生。

（1）只要端口使能STP，则配置BPDU就会按照Hello Time定时器规定的时间间隔从指定端口发出。

（2）当根端口收到配置BPDU时，根端口所在的设备会向自己的每一个指定端口复制一份配置BPDU。

（3）当指定端口收到比自己差的配置BPDU时，会立刻向下游设备发送自己的配置BPDU。

配置BPDU的长度至少要35个字节，包含了桥ID、路径开销和端口ID等参数，如图8-8所示。只有当“发送者BID”或“发送端口PID”两个字段中至少有一个和本桥接收端口不同，所收到的这个BPDU报文才会被处理，否则丢弃。这样避免了处理和本端口信息一致的BPDU报文。

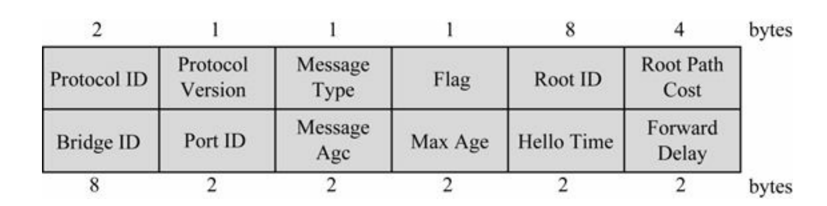


图8-8 STP BPDU报文格式

图8-8中所示的配置BPDU报文格式中的各部分说明如表8-1所示。其中Flags（标志）字段是一个字节长，但在STP配置BPDU中只用了最高位和最低位两个比特位，如图8-9所示，字段中的这两位说明参见表8-1中的Flags字段说明。

表8-1 配置BPDU报文基本格式

字段	字节数	说明
Protocol Identifier（协议 ID）	2	总是为 0
Protocol Version（协议版本）	1	总是为 0
Message Type（消息类型）	1	指示当前 BPDU 消息类型：0x00 为配置 BPDU，0x80 为 TCN BPDU
Flags（标志）	1	最低位=TC（Topology Change，拓扑变化）标志，最高位=TCA（Topology Change Acknowledgment，拓扑变化确认）标志
Root Identifier（根 ID）	8	指示当前根桥的 BID（即“**根 ID**”），由 2 字节的桥优先级和 6 字节 MAC 地址构成
Root Path Cost（根路径开销）	4	指示发送该 BPDU 报文的端口累计到根桥的开销
Bridge Identifier（桥 ID）	8	指示发送该 BPDU 报文的交换设备的 BID（即“**发送者 BID**”），也是由 2 字节的桥优先级和 6 字节 MAC 地址构成
Port Identifier（端口 ID）	2	指示发送该 BPDU 报文的端口 ID，即“**发送端口 ID**”
Message Age（消息生存时间）	2	指示该 BPDU 报文的生存时间，即端口保存 BPDU 的最长时间，过期后将删除，要在这个时间内转发才有效。如果配置 BPDU 是直接来自根桥的，则 Message Age 为 0，如果是其他桥转发的，则 Message Age 是从根桥发送到当前桥接收到 BPDU 的总时间，包括传输延时等。实际实现中，配置 BPDU 报文经过一个桥，Message Age 增加 1

（续表）

字段	字节数	说明
Max Age（最大生存时间）	2	指示配置 BPDU 消息的最大生存时间，也即老化时间
Hello Time（Hello 消息定时器）	2	指示发送两个相邻 BPDU 的时间间隔
Forward Delay（转发延时）	2	指示控制 Listening 和 Learning 状态的持续时间，表示在拓扑结构改变后，交换机在发送数据包前维持在监听和学习状态的时间

Bit7	Bit6	Bit5	Bit4	Bit3	Bit2	Bit1	Bit0
TCA	Reserved						TC

图8-9 STP BPDU中的Flags字段结构

2．TCN BPDU

TCN BPDU是指在下游拓扑发生变化时向上游发送拓扑变化通知，直到根桥。TCN BPDU在如下两种情况下会产生。

（1）端口状态变为Forwarding状态，且该设备上至少有一个指定端口。

（2）指定端口在收到TCN BPDU后向根桥复制TCN BPDU。

TCN BPDU中的内容比较简单，只有表 8-1中列出的前 3个字段：协议 ID、协议版本和消息类型，长度只有**4**个字节 ，且“消息类型”字段是固定值0x80。

3．BPDU优先级

前面说了，当桥收到其他桥发来的配置BPDU时会视情况对自己的配置BPDU进行更新，那么当桥的多个端口收到多个不同的配置BPDU时该以哪个为准呢？这就要使用配置BPDU的优先级来区分了。即当同一桥收到了多个不同的配置BPDU时，优先级高的BPDU则采用，其他的将被丢弃。假定有两条配置BPDU X和Y，则它们的比较顺序如下。

（1）如果X的根桥 ID小于Y的根桥 ID，则X优于Y。

（2）如果X和Y的根桥ID相同，但X的根路径开销小于Y，则X优于Y。

（3）如果X和Y的根桥ID，以及路径开销都相同，但X的桥ID小于Y，则X优于Y；

（4）如果X和Y的根桥ID、根路径开销以及桥ID都相同，但X的端口ID小于Y，则X优于Y。

8.1.5 STP的不足之处

STP协议虽然能够解决环路问题，但是由于网络拓扑收敛慢，影响了用户通信质量。因为在STP协议中任何端口要从Blocking（阻塞）状态转换到Forwarding（转发）状态必须经过两倍转发延时（包括由监听状态到学习状态的等待时间和由学习状态到转发状态的等待时间），至少30s时间。如果网络中的拓扑结构频繁变化，网络也会随之频繁失去连通性，从而导致用户通信频繁中断，这是用户无法忍受的。

STP的不足主要体现在以下几个方面。

（1）首先，STP没有细致区分端口状态和端口角色，不利于初学者学习及部署。

在STP协议中划分了5种端口状态，然而其中的Listening、Learning和Blocking这三种状态并没有实质上的区别，都不转发用户流量。另外，从使用和配置角度来讲，端口之间最本质的区别并不在于端口状态，而是在于端口扮演的角色。而在STP中，根端口和指定端口既可能都处于Listening状态，又可能都处于Forwarding状态，没有一个很好的体现。

（2）其次，STP协议采用的是被动算法，依赖定时器（如转发延时定时器）等待的方式判断拓扑变化，所以收敛速度慢。

（3）最后，STP 协议中的算法规定在稳定的拓扑中只有根桥才能主动发出配置BPDU 报文，而其他桥设备只能被动地进行转发，直到传遍整个 STP 网络。这也是导致拓扑收敛慢的主要原因之一。

正因STP有以上这些不足，IEEE于2001年发布的802.1W标准定义了RSTP（Rapid Spanning-Tree Protocol，快速生成树协议）。该协议基于STP协议，在绝大多数方面都是直接继承，但也针对STP协议的许多不足进行了比较多的修改和补充。

8.2 STP拓扑计算原理深入剖析

STP协议拓扑结构生成树的计算过程要区分初始化阶段和拓扑结构稳定后这两个阶段。在本章前面8.1.2节已介绍了STP中的根桥、根端口和指定端口的选举规则。本节要通过具体的示例再次深入剖析在初化阶段STP的这三个要素的选举原理，以及在拓扑稳定阶段，因拓扑发生变化而引起的生成树拓扑改变原理。

8.2.1 生成树初始化阶段的角色选举

网络中所有的桥设备在使能STP协议后，每一个桥设备都认为自己是根桥。此时每台设备仅仅收发配置BPDU，而不转发用户流量，所有的端口都处于Listening状态。所有桥设备通过交换配置BPDU后才进行根桥、根端口和指定端口的选举工作。

1．根桥的选举

“根桥的选举”就是在交换网络中所有运行 STP 协议的交换机上选举出一个唯一的根桥。“根桥”是STP生成树的最顶端交换设备，是STP生成树的“树根”。根桥的选举依据是各桥的配置BPDU报文中BID（桥ID）字段值，BID字段值最小的交换机将成为根桥。而桥配置BPDU报文中BID字段共有8个字节，即2个字节的桥优先级和6个字节的桥背板MAC，其中桥优先级的取值范围是 0～65 535，缺省值是 32 768。在进行BID比较时，先比较桥优先级，优先级值小的为根桥；当桥优先级值相等时，再比较桥的背板MAC地址，MAC地址小的为根桥。

在初始化过程中，根桥的选举要经历两个主要过程：一是每桥上确定自己的配置BPDU；二是在整个交换网络中通过各桥自己发送的配置BPDU进行比较选举整个交换网络中的根桥。

（1）桥配置BPDU的确定。一开始每个桥都认为自己是根桥，所以在每个端口所发出的配置BPDU报文中，“根ID”字段都是用各自的BID，“根路径开销”字段值均为0， “发送者BID”字段是自己的BID，“发送端口PID”字段是发送该BPDU端口的端口ID。

每个桥在向外发送自己的配置 BPDU 的同时也会收到其他桥发送的配置 BPDU。但桥端口并不会对收到的所有配置 BPDU都用来更新自己的配置 BPDU，而是先会进行配置BPDU优先级比较。当端口收到的配置BPDU比本端口的配置BPDU的优先级低时，将丢弃所收到的这个配置 BPDU，仍保留自己原来的配置BPDU，否则桥将收到的配置 BPDU 作为该端口的配置 BPDU。然后，桥再将自己所有端口的配置 BPDU进行比较，选出最优的BPDU作为本桥的配置BPDU。有关BPDU优先级的比较参见本章8.1.4节。

（2）根桥的确定。每个桥的最优配置BPDU确定后，以后各桥间交换的配置BPDU都是各自最优的配置 BPDU 了。如图 8-10 所示，用{}标注的四元组表示了由根桥 BID （图中以 S1_MAC 和 S2_MAC 代表两台设备的 BID）、累计根路径开销、发送者 BID （SBID）、发送端口PID构成的有序组。配置BPDU会按照Hello Timer规定的时间间隔来发送，缺省的时间是2s。

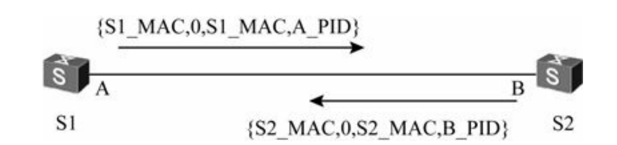


图8-10 初始信息交互过程示例

一旦某个端口收到比自己优的配置BPDU报文，此端口就提取该配置BPDU报文中的某些信息更新自己的信息。该端口存储更新后的配置BPDU报文后，立即停止发送自己的配置BPDU报文。在图中，如果S2的端口B由于接收到了来自S1的更好的配置BPDU，从而认为此时S1是根桥，然后S2的其他端口再发送BPDU的时候，在根桥ID字段里面填充的就是 S1_BID 了。此过程不断交互进行，直到所有交换设备的所有端口都认为根桥是相同的，说明根桥已经选择完毕。

在如图8-11所示的交换网络中列出了S1、S2和S3的桥优先级和桥MAC地址。通过比较发现三台交换机的桥优先级都一样，均为缺省的 32 768，这时就要进一步比较各交换机的MAC地址，通过比较可以发现S1的MAC地址最小，所以最终S1将选举作为根桥。

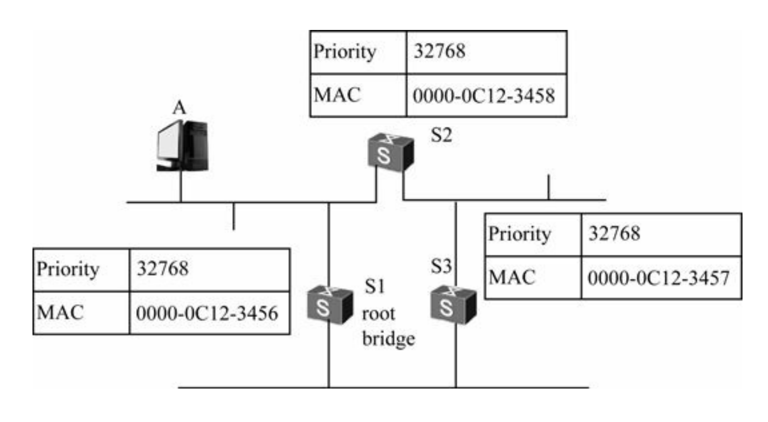


图8-11 根桥选举示例

2．根端口的选举

“根端口的选举”就是在所有非根桥上的不同端口之间选举出一个到根桥最近的端口。当然这个“最近”的衡量标准不是根据到达根桥所经过的桥数，而是根据端口到根桥的累计根路径开销最小来判定的。实质上

是非根桥上接收到最优配置BPDU的那个端口即为根端口。每个非根桥设备都要选择一个根端口，根端口对于一个设备来说有且只有一个 。

累计根路径开销的计算方法是累加从端口到达根桥所在路径的各端口（除根桥上的指定端口外 ）的各段链路的路径开销值（也称链路开销值）。这里要特别注意的是，同一交换机上不同端口之间的路径开销值为 **0** 。如果同一桥上有两个以上的端口计算得到的累计根路径开销相同，那么选择收到发送者BID最小的那个端口作为根端口。

在如图8-12所示的交换网络中，S1为根桥，这时就需要选举S2和S3非根桥的根端口。S2到达根桥S1有两条路径：一条是通过port5端口直接到达S1的port1端口，其累计根路径开销很容易得出，就是port5端口自身的路径开销值，即图中标的是19。另一条是从port6端口出发，经过S3的port3和port4端口，到达根桥S1的port2端口，其累计根路径开销值就是port6、port3和port4端口的路径开销值之和。从图中的标注可以知道，port6端口的路径开销值为也为19，但因为port3到port4端口在同一交换机S3上，所以port3到port4端口的路径开销值为0，port4到S1的port2端口的路径开销值也为19，这样port6端口累计根路径开销值就是19+0+19=38，很明显高于port5端口的累计根路开销值19，所以port5端口最终选举为S2的根端口。用同样的方法可以得出S3桥上的根端口为port4。

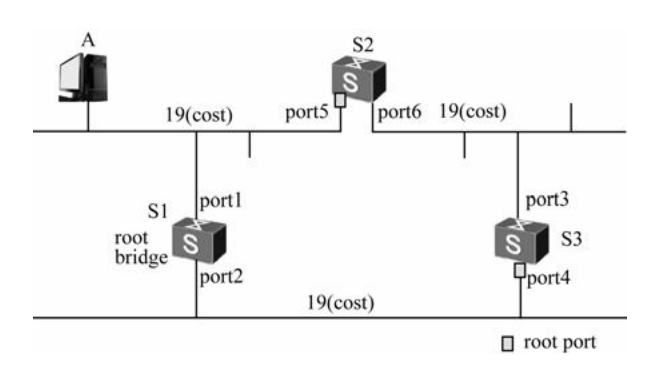


图8-12 根端口选举示例

3．指定端口的选举

“指定端口的选举”是在每一个物理网段的不同端口之间选举出一个指定端口。“指定端口”与前面所说的“根端口”相对，它可以理解为离下游设备最近的端口，是本物理网段（这里的“网段”是指一个交换机端口所连接的所有设备）中唯一可以接收下游设备数据的端口。它是依次根据以下三项条件来判定的。

（1）某网段到根桥的路径开销最小。

（2）接收数据时发送方（也就是链路对端的桥）的桥ID最小

（3）发送方端口ID最小（端口ID有16位，它是由8位端口优先级和8位端口编 号组成的，其中端口优先级的取值范围是0～240，缺省值是128，可以修改，但必须是16的倍数）。

如图8-10所示，假定S1的MAC地址小于S2的MAC地址，则S1为根桥。根据上面的第一项指定端口判定原则可以得出S1的端口A会成为指定端口。在一个物理网段上拥有指定端口的设备被称作该网段的指定桥，由此可以得出图8-10所示的S1-S2间网段的指定桥是S1。

网络收敛后，只有指定端口和根端口可以处于转发状态。其他端口都是Blocking状态，不转发用户流量。根桥的所有端口都是指定端口（除根桥物理上存在环路）。

现在再来看如图 8-13 所示交换网络中指定端口的选举。S1 为根桥，这样很容易根据前面列出的指定端口判定条件中的第一项得出在S2-S1网段，以及S3-S1网段中的指定端口分别为S1的port1和port2端口。而在S3-S2网段中，由于S3和S2桥到达根桥的路径开销均为 19，所以这里要比较前面提到的第二项条件，即发

送方的桥 ID（即图中标识的SBID）大小了。S3的port3的发送方的桥ID为32768.000-0C12-3457，而S2的port6的发送方的桥ID为32768.000-0C12-3458，经过比较发现S3的port3的发送方的桥ID更小，所以最终选举为S3-S2网段的指定端口。这样一来就可确定port6端口为阻塞端口了。

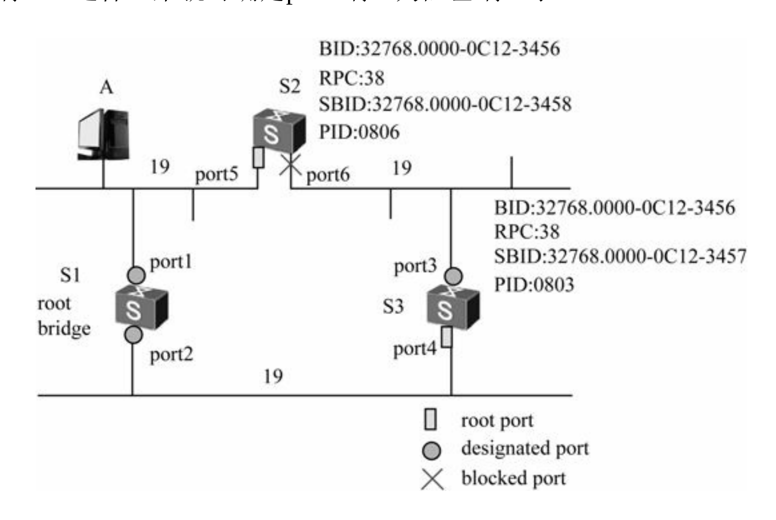


图8-13 指定端口选举示例

8.2.2 拓扑发生变化后的角色选举

拓扑稳定后，根桥仍然按照Hello定时器规定的时间间隔发送配置BPDU报文，非根桥设备从根端口收到配置BPDU报文，通过指定端口转发。如果接收到优先级比自己高的配置BPDU，则非根桥设备会根据收到的配置BPDU中携带的信息更新自己相应的端口存储的配置BPDU信息。

在网络拓扑结构发生变化后，下游设备会不间断地向上游设备发送 TCN BPDU报文。上游设备在收到下游设备发来的TCN BPDU报文后，只有指定端口处理TCN BPDU报文。其他端口也有可能收到TCN BPDU报文，但不会处理。上游设备会使用Flags字段中TCA（拓扑变化确认）标志位置1的配置BPDU报文发送给下游设备，告知下游设 备停止发送TCN BPDU报文。与此同时，上游设备复制一份TCN BPDU报文，向根桥方向发送。当根桥收到TCN BPDU报文后，根桥会使用 Flags字段中TC（拓扑变化）标志位置1的配置BPDU报文向对应下游设备回送，通知下游设备直接删除发生故障的端口的MAC地址表项。

由此可以看出，在发生拓扑变化时，下游设备使用TCN BPDU报文向上游设备通知，但上游设备使用的是TC位，或者TCA位置 1的配置BPDU，而不是TCN BPDU通知下游设备。即TCN BPDU报文用来向上游设备乃至根桥通知拓扑变化；TCA标志位置 1的配置BPDU报文主要是上游设备用来告知下游设备已经知道拓扑变化，通知下游设备停止发送TCN BPDU报文；TC标志位置 1的配置BPDU报文主要是上游设备用来告知下游设备拓扑发生变化，要求下游设备直接删除有故障的端口的 MAC 地址表项，从而达到快速收敛的目的。

下面以图8-13为例说明根桥、根桥的指定端口分别发生故障时，网络拓扑如何收敛。

当根桥发生故障时，设备S2和设备S3之间将重新选举根桥。设备S2和设备S3之间根据交互的配置BPDU报文，选出新的根桥S3，如图8-14所示。再假设根桥指定port1端口发生故障时，S2和S3通过交互配置BPDU报文将port6选举为根端口，如图8-15所示。同时，port6变为forwarding状态后，会向外发送TCN报文，根桥收到TCN报文后向其他设备发送TC报文，通知其他设备直接删除MAC表项。

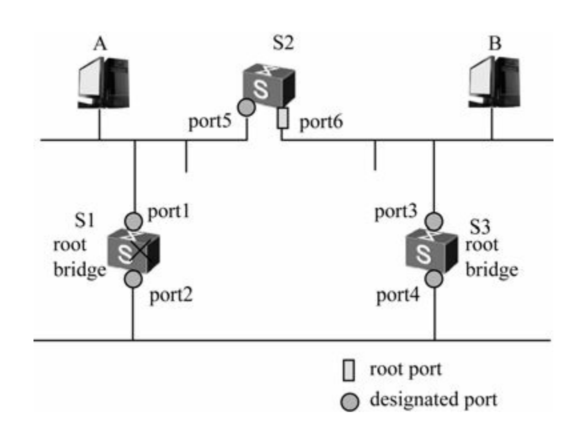


图8-14 根桥发生故障时重新选举新的根桥的示例

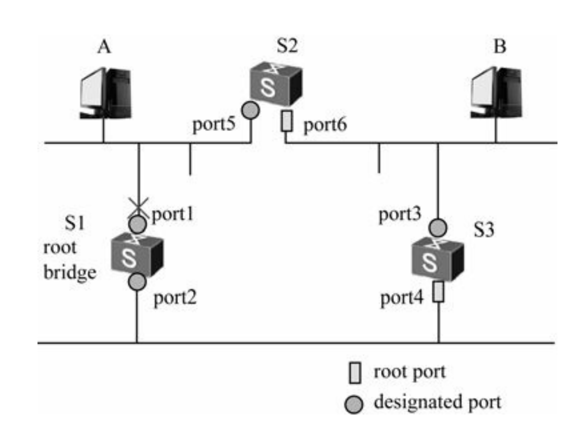


图8-15 指定端口发生故障时重新选举新的根端口的示例

8.3 RSTP对**STP**的改进

继 IEEE 802.1D定义了STP标准后，IEEE又推出了 802.1w这个草案作为 802.1D的补充，并定义了RSTP标准。在新版本的802.1D（2004）中已经接纳了RSTP标准，取代了原来的STP。RSTP保留了STP的大部分算法和计时器，只在一些细节上做了改进。但这些改进相当关键，极大地提升了STP的性能，使其能满足如今低延时高可靠性的网络要求。本章后面将要介绍的MSTP，在单个实例中的算法和RSTP几乎一模一样，所以可以说从STP发展到RSTP的这套算法，是整个生成树协议的精髓。

根据8.1.5节介绍的STP协议的不足，RSTP协议删除了那3种区分不明显的3种端口状态，另外新增加了两种端口角色，并且解除了端口属性中端口状态和端口角色的关联，使得可以更加精确地描述端口，从而使得初学者更易学习协议，同时也加快了拓扑收敛。

8.3.1 新增三种端口角色

RSTP 的端口角色共有 5 种：根端口、指定端口、Alternate（替代）端口、Backup （备份）端口和Edge（边缘）。前4种端口角色如图8-16所示。

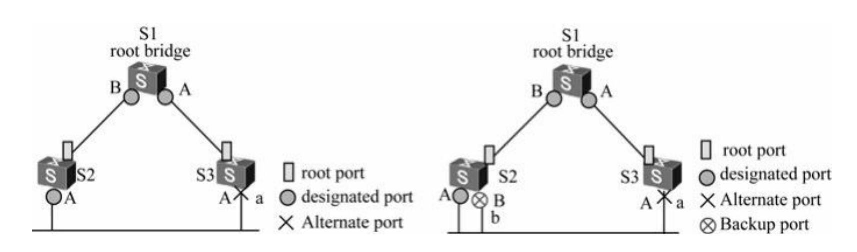


图8-16 RSTP的4种端口角色

当“根端口”或“指定端口”失效时，“替代端口”或“备份端口”就会无延时地进入转发状态，提高了收敛效率。Edge端口是管理员根据实际需要配置的一种指定端口，用以连接PC或不需要运行STP的下游交换机。管理员需要保证该端口下游不存在环路， Edge端口能够直接进入Forwarding状态。

在RSTP中，“根端口”和“指定端口”的作用与STP协议中对应的端口角色一样，而Alternate端口和Backup端口从配置BPDU报文发送角度来看：Alternate端口就是由于学习到其他桥发送的配置BPDU报文而阻塞的端口，Backup端口就是由于学习到自己发送的配置BPDU报文而阻塞的端口。从用户流量角度来看：Alternate端口提供了从指定桥到根桥的另一条可切换路径，作为根端口的备份端口；而 Backup 端口是作为指定端口的备份，提供另一条从根桥到相应网段的备份通路。

给一个RSTP域内所有端口分配角色的过程就是整个拓扑收敛的过程，远比STP协议中要同时顾及端口角色及端口状态的效率高。

8.3.2 重新划分端口状态

在端口状态上，RSTP也对STP做了比较大的改进，把STP的5种状态缩减为3种。并且是根据端口是否转发用户流量和学习MAC地址来进行划分的，具体如下。

（1）如果不转发用户流量也不学习MAC地址，那么端口状态就是Discarding（丢弃）状态。

（2）如果不转发用户流量但是学习MAC地址，那么端口状态就是Learning（学习）状态。

（3）如果既转发用户流量又学习MAC地址，那么端口状态就是Forwarding（转发）状态。

表 8-2 是对 STP 中的端口状态与 RSTP 的端口状态的比较。从中可以看出，RSTP中的端口状态和端口角色是没有必然关联的。由以上可以看出，RSTP 只有 3 种端口状 态：Discarding（丢弃）、Leaning（学习）和Forwarding（转发），它把STP中的Blocking （阻塞）、Listening（监听）和Disabled（禁用）三种状态统一用一种状态——Discarding （丢弃）替代。这样的好处就是一个端口从初始状态转变为转发状态只需要一个转发延时周期时间，也就是从学习状态到转发状态所需等待的时间。在活跃拓扑中，只有“学习”和“转发”这两种状态的端口。

表8-2 STP与RSTP端口状态比较

STP 端口状态	RSTP 端口状态	对应的端口角色	是否发送 BPDU	是否学习 MAC 地址	是否 发送数据
Forwarding	Forwarding	包括根端口、指定端口	是	是	是
Learning	Learning	包括根端口、指定端口	是	是	否
Listening	Discarding	包括根端口、指定端口	否	否	否
Blocking	Discarding	包括 Alternate 端口、Backup 端口	否	否	否
Disabled	Discarding	包括 Disable	否	否	否

8.3.3 BPDU的改变

RSTP协议与STP协议一样也是使用BPDU消息格式进行各桥间的拓扑信息交互的，但是它只有配置BPDU，没有TCN BPDU，且RSTP的配置BPDU称为RST BPDU。RSTP在BPDU方面的改变主要体现在BPDU格式，在发生拓扑改变时BPDU的使用，以及对BPDU的处理3个方面，下面分别予以具体介绍。

1．BPDU格式上的改变

在BPDU格式上，RSTP的RST BPDU与STP的配置BPDU没做什么重大修改，只是对STP配置BPDU中的Flag（标志）字段进行了填充，使从RST BPDU中就可以看出对应端口的端口角色，另外就是在 BPDU类型值上做了改变。具体表现在以下两个字段。

（1）Type字段：RST BPDU类型不再是 0，而是 2，所以运行STP的设备收到RSTP的RST BPDU时会丢弃。

（2）Flag字段：在RST BPDU中使用了在 STP配置BPDU中该字段保留的中间 6位（最高位仍为TCA，最低位仍为TC），如图8-17所示。中间6位的作用如下。

① Agreement：确认标志位，位于Bit6，当该位置1时，表示该BPDU报文为快速收敛机制中的Agreement报文，是对所收到的 Proposal BPDU（此时Bit1位置 1）的提议进行确认。RSTP中定义了Proposal/Agreement机制（提议/确认机制，即P/A机制），可使指定端口通过与对端端口进行一次握手即可快速进入转发状态，其中不需要任何定时器。

②Forwarding：转发状态标志位，位于Bit5，当该位置1后表示发送该BPDU报文的端口处于Forwarding状态。

③Learning：学习状态标志位，位于Bit4，当该位置1后表示发送该BPDU报文的端口处于Learning状态。

④ Port role：端口角色标志位，位于Bit3和Bit2共两位，取值为 00时表示发送该 BPDU的端口的角色未知，为01时表示该端口为Alternate端口或Backup端口，为10时表示该端口为根端口，为11时表示该端口为指定端口。

⑤ Proposal：提议标志位，位于Bit1，当该位置1时表示该BPDU报文为快速收敛机制中的Proposal报文。对端在收到该报文后，如果同意，则需要发送Bit6位置1的确认报文。

Bit7	Bit6	Bit5	Bit4	Bit3	Bit2	Bit1	Bit0
TCA	Agreement	Forwarding	Learning	Port role		Proposal	TC

图8-17 RSTP BPDU中的Flags字段结构

从以上介绍可以得知，RSTP的Flags字段增加了端口属性和状态，其中Bit1和Bit6两个字段用于点到点链路端口快速收敛中的消息报文。常见的几种Flags需要记住：2c，即 00101100，表示发送 BPDU 的端口状态为转发状态，端口角色为指定端口；0e，即00001110，表示是由指定端口发送的提议 BPDU 报文；6c，即 01101100，表示是由处于转发状态的指定端口发送的确认BPDU报文；2d，即00101101，表示是由处于转发状态的指定端口发送的拓扑更改BPDU报文。

注意

运行STP的设备会丢弃收到的RST BPDU，目前RSTP交换机都提供STP兼容模 式，运行在STP兼容模式的端口会发送和接收配置BPDU，表现的特性也和STP类似。除了配置BPDU外，RSTP同样有TCN BPDU，类型值也为 0x80。

2．拓扑变化时BPDU的使用变化

通过前面的学习我们知道，在 STP 协议中只要有端口变为 Forwarding 状态，或从Forwarding状态转变

到Blocking状态均会触发拓扑改变处理过程。在发生拓扑变化时，下游设备会不间断地向上游设备发送TCN BPDU报文。上游设备在收到下游设备发来的TCN BPDU报文后，使用Flags字段中TCA标志位置 1的配置BPDU报文发送给对应的下游设备，告知下游设备停止发送TCN BPDU报文。与此同时，上游设备复制一份TCN BPDU报文，向根桥方向发送。当根桥收到TCN BPDU报文后，根桥又使用Flags字段中TC标志位置1的配置BPDU报文向对应的下游设备回送，通知它们直接删除发生故障的端口的MAC地址表项。整个过程同时使用了TCN BPDU和配置BPDU。

在 RSTP 协议中检测拓扑是否发生变化只有一个标准：一个非边缘端口迁移到Forwarding状态。一旦检测到拓扑发生变化，将进行如下处理。

（1）为本设备上的所有非边缘指定端口启动一个 TC While定时器，该定时器值是Hello定时器的两倍。在这个时间内，清空状态发生变化的端口上学习到的MAC地址。同时，由这些端口向外发送TC位置 1的RST BPDU。一旦TC While定时器超时，则停止发送RST BPDU。

（2）其他交换设备接收到TC位置1的RST BPDU后，清空所有端口学习到的MAC地址，除了收到该RST BPDU报文的端口。然后也为自己所有的非边缘指定端口和根端口启动TC While定时器，重复上述过程。

如此，网络中就会产生RST BPDU的泛洪。由此可见，在RSTP协议中不再使用TCN BPDU，而是发送TC位置 1的RST BPDU，并通过泛洪的方式快速通知整个网络。不需 要依次向上发送TCN BPDU至根桥，当其他桥收到TC位置1的RST BPDU之后，也不再需要等待由根桥向下发送的TC位置 1的 RST BPDU，直接清除端口学习到的MAC地址，重新学习，实现网络的快速收敛。

3．配置BPDU处理上的变化

RSTP在配置BPDU处理上的变化主要体现在以下几个方面。

（1）拓扑稳定后，配置BPDU报文的发送方式。在STP协议中，当拓扑稳定后，只能由根桥按照 Hello定时器规定的时间间隔定期发送配置 BPDU，其他非根桥设备只能在收到上游设备发送过来的配置BPDU后才会触发发出配置BPDU。此方式使得STP协议计算复杂且缓慢。RSTP 协议对此进行了改进，即在拓扑稳定后无论非根桥设备是否接收到根桥传来的RST BPDU报文，非根桥设备仍然按照Hello定时器规定的时间间隔定期发送配置BPDU。即在RSTP中，各桥的配置BPDU发送行为完全是由每台桥设备自主进行。

（2）更短的BPDU超时计时。在RSTP协议中规定，如果一个端口连续3倍Hello定时器时间内没有收到上游设备发送过来的RST BPDU，那么该设备认为与此邻居之间的协商失败。而不是像 STP协议规定的那样需要先等待一个Max Age（最大生存时间）。

（3）处理次等BPDU。在STP中指定端口在收到inferior（次优）BPDU会马上把端口保存的更优的BPDU发送出去，但对非指定端口不会做同样处理。而在RSTP中不管是否指定端口，收到次优RST BPDU都会马上发送本地更优的RST BPDU给对端，以使对端口快速更新自己的RST BPDU。具体为，当一个端口收到上游的指定桥发来的RST BPDU报文时，该端口会将自身存储的RST BPDU与收到的RST BPDU进行比较。如果该端口存储的RST BPDU的优先级高于收到的RST BPDU，那么该端口会直接丢弃收到的RST BPDU，立即以自身存储的RST BPDU进行响应。当上游设备收到下游设备响应的RST BPDU后，上游设备会根据收到的RST BPDU报文中相应的字段立即更新自己存储的RST BPDU。由此可见，RSTP处理次等BPDU报文不再像STP那样依赖于任何定时器通过超时解决拓扑收敛，从而加快了拓扑收敛。

在如图8-18中，假设桥优先级S3<S2<S1，各段链路的开销售值在图中已进行了标注。正常情况下，S1为根桥，S2的port1端口为S2的根端口，S2的port2端口为S2的指定端口，S3的port2端口为S3的根端口，S3的port1端口为阻塞端口。因为S3经过S2到达根桥S1的开销更小，所以S3会从port2端口发送数据，而不会从port1端口发送数据。

如果现在S1、S2中的链路down了，如图8-19所示。在STP协议中，一开始S2会认为自己是根桥（因为此时S2误认为S1不存在了），并发送配置BPDU。但S3的port2不会立即以更优的BPDU响应S2，直到Max age（缺省为 20s）过期，即S3的port2端口上保存的原BPDU超时，S3的port2端口才发送新的以S1为根（因为此时S3与S1仍然保持着连接，S1的优先级更高，所以S3认为S1仍为该交换网络的根桥）的BPDU。S2接收到这个BPDU后，承认S1为根桥（原因是S1的优先级高于S2），修改自己的桥角色。此时S2的port2端口状态会发生一系列的变化，需经过两倍转发延时——30s才进入转发状态。这样一来就一共经历50s的时间。

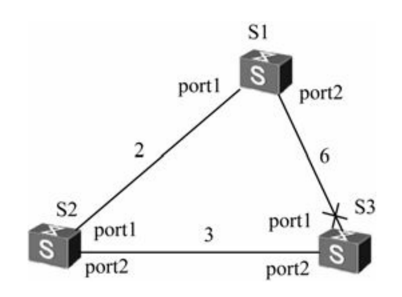


图8-18 对次优BPDU处理的示例图一

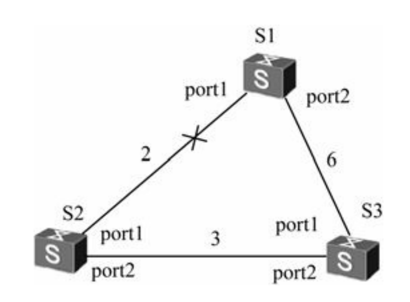


图8-19 对次优BPDU处理的示例图二

如果是在RSTP中，S3的port2端口收到S2发来的次BPDU后会马上发送本端口的更优BPDU，无需等待Max Age时间。因为此时S2的 port2与S3的 port2端口是点对点连接，所以这两个端口都能快速地进行状态迁移，即由 Discarding 状态迁移到Forwarding，实现拓扑瞬间收敛，无需等待两倍的转发延时。

8.3.4 更加快速的**P/A**收敛机制

在RSTP中，为了实现更加快速的拓扑收敛，主要采用了Proposal/Agreement（提议/确认，P/A）机制。

通过本章前面的学习已经知道，当一个端口被选举成为指定端口之后，在STP协议中该端口至少要等待两个Forward Delay的时间（由Listening状态到Learning状态，再从Learning状态到Forwarding状态）才会迁移到Forwarding状态，发送数据。这种保守的设计可以保证不产生环路，但显然不够聪明，RSTP 对此做了一系列改进。在 RSTP协议中规定，一个端口被选举为指定端口后，会先进入 Discarding 状态，再通过Proposal/Agreement 机制快速进入 Forwarding 状态。但这种机制必须在点到点链路 （某端口在所属的共享以太网上的对端只有一台设备，这样的以太网被认为是点到点链路）上使用 。

1．P/A机制工作原理

P/A机制即Proposal/Agreement机制，其目的是使一个指定端口尽快进入Forwarding状态。如图8-20所示，P/A协商过程的完成根据以下几个端口变量的协商，也是P/A机制的具体工作原理。

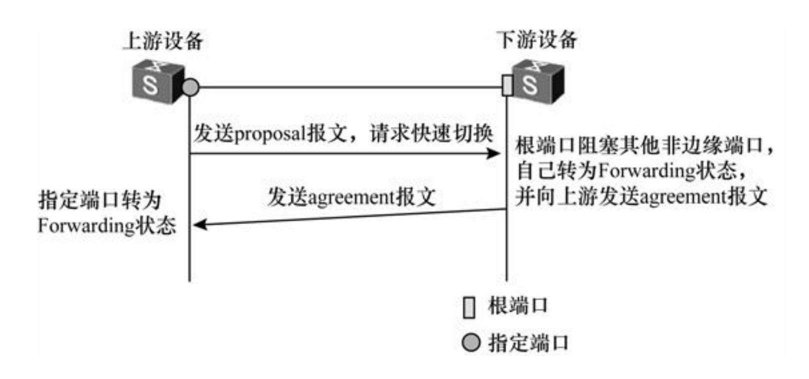


图8-20 P/A机制工作原理示意图

（1）proposing（提议请求）：当一个指定端口处于 Discarding 或 Learning 状态时， proposing变量置位，并向下游桥设备传递Flags字段Proposal标志位置1的RST BPDU （表示此BPDU为Proposal RST BPDU报文），请求快速切换到Forwarding状态。

（2）proposed（提议采纳）：当对端的根端口收到以上Proposal标志位置1的RST BPDU时，proposed变量置位。该变量指示本网段上的指定端口希望尽快进入Forwarding状态。

（3）sync（同步请求）：当根端口的 proposed 变量置位后会依次为本桥上的其他端口使sync变量置位，使所有非边缘端口都进入Discarding状态，准备重新同步。

（4）synced（同步完成）：当端口进入到Discarding状态后会将自己的synced变量置位，包括本桥上的其他所有Alternate端口、Backup端口和边缘端口，实施同步操作。此时，根端口监视其他端口的synced变量置位情况，当所有其他端口的synced变量全被置位，则根端口最后也会将自己的synced变量置位，表示本桥上已正式完成同步操作，向上游设备传回Agreement标志位置1的RST BPDU（表示此BPDU为Agreement RST BPDU报文）。

（5）agreed（提议确认）：当原来想要进入转发状态的上游设备指定端口收到对端根端口发来的一个Agreement RST BPDU时，则此指定端口的 agreed变量被置位。Agreed变量一旦被置位，则该指定端口马上转入Forwarding状态。

2．P/A机制解析示例

如图8-21所示，根桥S1和S2之间新添加了一条链路。在当前状态下，S2的另外 几个端口p2是Alternate端口，p3是指定端口且处于Forwarding状态，p4是边缘端口。新链路连接成功后，P/A机制协商过程如下。

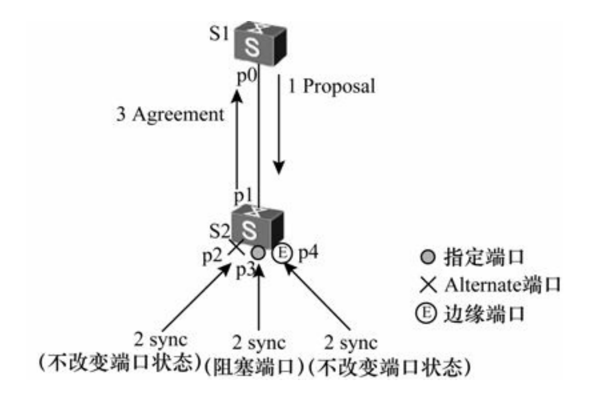


图8-21 P/A机制解析示例

（1）p0和p1两个端口马上成为指定端口，发送RST BPDU。

（2）S2 的 p1 端口在收到更优的 RST BPDU后马上意识到自己将成为根端口，而不是指定端口，于是停止发送RST BPDU。

（3）当S1的p0端口处于Discarding状态（这是所有端口的初始状态）时向对端的S2发送 proposal标志位置 1的RST BPDU。

（4）S2收到根桥发送来的携带proposal标志位的RST BPDU后开始将自己的所有端口的sync变量置位，进入Discarding 状态。因为此时p2 端口已经阻塞，所以状态不变，p4端口是边缘端口不参与运算，所以只需要阻塞非边缘指定端口p3。

（5）在p2、p3、p4端口都进入Discarding状态之后，各处将自己的synced变量置位，然后根端口p1也将自己的synced变量置位，然后向S1返回Agreement标志位置1的响应RST BPDU。该RST BPDU携带和刚才根桥发过来的BPDU一样的信息，除了Agreement标志位置1之外（Proposal位清零）。

（6）当S1判断出所收到的Agreement RST BPDU是对刚刚发出的Proposal的响应后，端口p0马上进入Forwarding状态。

以上P/A过程可以向下游设备继续传递，也就是说不一定是根桥与非根桥之间，也可以是非根桥之间。

说明

事实上对于STP，指定端口的选择可以很快完成，主要的速度瓶颈：为了避免环 路，必须等待足够长的时间，使全网的端口状态全部确定，也就是说必须要等待至少一个Forward Delay所有端口才能进行转发。而RSTP的主要目的就是消除这个瓶颈，通过阻塞自己的非根端口来保证不会出现环路。而使用 P/A 机制加快了上游端口转到Forwarding状态的速度。

但P/A机制要求两台交换设备之间链路必须是点对点的全双工模式。一旦P/A协商不成功，指定端口的选择就需要等待两个Forward Delay，协商过程与STP一样。

8.3.5 RSTP的其他收敛机制和与STP的互操作

1．其他收敛机制

在RSTP中，除了P/A机制外，还有以下两种机制也可帮助实现拓扑的快速收敛。

（1）根端口快速切换机制。如果网络中一个根端口失效，那么网络中最优的 Alternate端口将立即成为根端口，并进入 Forwarding 状态。在点到点以太网链路上，根端口总能快速迁移到Forwarding状态。

（2）边缘端口的引入。在RSTP里面，如果某一个指定端口位于整个网络的边缘，即不再与其他交换设备连接，而是直接与终端设备直连，这种端口叫做边缘端口。

边缘端口不接收处理配置BPDU，不参与RSTP运算，可以由Disable状态直接转到Forwarding状态，且不经历任何时延。但是一旦边缘端口收到配置BPDU，就丧失了边缘端口属性，成为普通RSTP端口，并重新进行生成树计算，从而引起网络振荡。

2．RSTP与STP的互操作

RSTP可以和STP互操作，但是此时会丧失快速收敛等RSTP优势。当一个网段里既有运行 STP 的交换设备，又有运行 RSTP 的交换设备时，STP 交换设备会忽略 RST BPDU，而运行 RSTP 的交换设备在某端口上接收到运行 STP 的交换设备发出的配置BPDU时，会在两个Hello Time时间之后把自己的端口转换到STP工作模式，发送配置BPDU。这样，就实现了互操作。

在华为技术有限公司的数据通信设备上可以配置运行STP的交换设备，被撤离网络后，运行RSTP的交换设备又可迁移回到RSTP工作模式。

8.4 STP/RSTP配置

虽然RSTP对STP有了重大改进，但在配置方法方面总体来说区别不大（主要区别体现在收敛参数方面的配置），故本节一起介绍STP 和 RSTP 协议的配置与管理方法。先来看一下华为S系列交换机上STP和RSTP的一些主要配置任务。

8.4.1 STP/RSTP配置任务及缺省配置

STP/RSTP可阻塞二层网络中的冗余链路，将网络修剪成树状，达到消除环路的目的。

（1）为了消除设备间的环路，可以配置STP/RSTP基本功能。

（2）为了加快设备的收敛速度，可以配置影响STP/RSTP拓扑收敛的参数。

（3）为了实现与其他厂商设备的互通，需要在华为公司运行STP/RSTP的设备上配 置合适的参数，以确保通信畅通。

（4）为了满足特殊场合的应用和功能扩展，还可配置RSTP拓扑收敛反馈机制，以及RSTP提供如表8-3所示的各种保护功能。

表8-3 RSTP保护功能

保护功能	场景	配置影响
BPDU保护	边缘端口在收到 BPDU 以后端口状态将变为非边缘端口，此时就会造成生成树的重新计算，如果攻击者伪造配置消息恶意攻击交换设备，就会引起网络振荡	交换设备上启动了 BPDU 保护功能后，如果边缘端口收到 RST BPDU，边缘端口将被 error-down，但是边缘端口属性不变，同时通知网管系统 被错误 down 掉的边缘端口只能由网络管理员手动恢复。如果用户需要被错误 down 掉的边缘端口可自动恢复，可通过配置使能端口自动恢复功能，并可设置延迟时间
防 TC-BPDU 报文攻击保护	交换设备在接收到拓扑变化报文后，会执行 MAC 地址表项和 ARP 表项的删除操作，如果频繁操作则会对 CPU 的冲击很大	启用防 TC-BPDU 报文攻击功能后，在单位时间内交换设备处理拓扑变化报文的次数可配置。如果在单位时间内交换设备在收到拓扑变化报文数量大于配置的阈值，那么设备只会处理阈值指定的次数。对于其他超出阈值的拓扑变化报文，定时器到期后设备只对其统一处理一次。这样可以避免频繁地删除 MAC 地址表项和 ARP 表项，从而达到保护设备的目的
Root 保护	由于维护人员的错误配置或网络中的恶意攻击，根桥收到优先级更高的 BPDU，会失去根桥的地位，重新进行生成树的计算，并且由于拓扑结构的变化，可能造成高速流量迁移到低速链路上，引起网络拥塞	对于启用 Root 保护功能的指定端口，其端口角色只能保持为指定端口。一旦启用 Root 保护功能的指定端口收到优先级更高的 RST BPDU 时，端口状态将进入 Discarding 状态，不再转发报文。在经过一段时间（通常为两倍的 Forward Delay），如果端口一直没有再收到优先级较高的 RST BPDU，端口会自动恢复到正常的 Forwarding 状态
环路保护	当出现链路拥塞或者单向链路故障，根端口和 Alternate 端口会被老化。根端口老化会导致系统重新选择根端口（而这有可能是错误的），Alternate 端口老化将迁移到 Forwarding 状态，这样会产生环路	在启动了环路保护功能后，如果根端口或 Alternate 端口长时间收不到来自上游的 RST BPDU 时，则向网管发出通知信息（如果是根端口则进入 Discarding 状态）。而阻塞端口则会一直保持在阻塞状态，不转发报文，从而不会在网络中形成环路。直到根端口收到 RST BPDU，端口状态才恢复正常到 Forwarding 状态

支持STP/RSTP的华为S系列交换机都有如表8-4所示的缺省配置，实际应用的配置可以基于缺省配置进行修改。

表8-4 STP/RSTP缺省配置

参数	缺省值
生成树协议工作模式	MSTP 模式
STP/RSTP 功能	全局 STP/RSTP 功能使能，接口的 STP/RSTP 功能也使能
交换设备的优先级	32768
端口的优先级	128
路径开销缺省值的计算方法	Dot1t，即 IEEE 802.1t 标准
Forward Delay Time	1 500 厘秒
Hello Time	200 厘秒
Max Age Time	2 000 厘秒

8.4.2 配置STP/RSTP基本功能

在以太网中，通过对交换设备配置STP/RSTP基本功能，将网络修剪成树状，达到消除环路的目的。下面先具体了解一下STP、RSTP基本功能的配置任务。

1．主要配置任务

STP/RSTP 基本功能配置包括 STP/RSTP 工作模式配置，根桥和备份桥配置，桥优先级配置，端口路径开销、端口优先级、STP或RSTP功能的启用等。当然其中大部分是为可选的配置任务，所以总体上配置还是很简单的。具体如下。

（1）配置STP/RSTP工作模式。华为S系列交换机支持STP、RSTP和MSTP三种生成树工作模式。在只运行STP的环形网络中可选择STP模式；在只运行RSTP的环形网络中可选择RSTP模式。其他情况，建议选择缺省情况MSTP模式。

（2）（可选）配置根桥和备份根桥。此为可选配置任务，因为缺省情况下，根桥和备份根桥是通过选举产生的。如果配置此项配置任务就相当于人工指定根桥和备份桥。但要注意，在同一交换机上只能选择配置根桥或者备份根桥，不能同时配置。在配置STP/RSTP过程中，建议手动配置根桥和备份根桥。

说明

在一棵生成树中，生效的根桥只有一个；在同一个网络中，当多个设备的桥优先 级相同时系统将选择 MAC 地址最小的设备作为根桥。可以在每棵生成树中指定多个备份根桥。当根桥出现故障或被关机时，备份根桥可以取代根桥成为指定生成树的根桥；但此时若配置了新的根桥，则备份根桥将不会成为根桥。如果配置了多个备份根桥，在当前根桥出现了故障时，则MAC地址最小的备份根桥将成为指定生成树的根桥。

（3）（可选）配置交换设备优先级。在一个运行STP/RSTP的网络中，有且仅有一个根桥，它是整棵生成树的逻辑中心。在进行根桥的选择时，一般会希望选择性能高、网络层次高的交换设备作为根桥。但是性能高、网络层次高的交换设备其优先级不一定高，因此需要配置优先级以保证该设备成为根桥。同时，对于网络中部分性能低、网络层次低的交换设备，不适合作为根桥设备，一般会配置其较低优先级以保证该设备不会成为根桥。但要注意的是，在配置交换设备的优先级数值时：数值越小，优先级越高 ，成为根桥的可能性越大。

（4）（可选）配置端口路径开销。路径开销是一个端口量，是STP/RSTP协议用于选择链路的参考值。端口路径开销值取值范围由路径开销计算方法决定。当确定路径开销计算方法后，端口所处链路的速率值越大，建议将该端口的路径开销值在指定范围内设置越小。

华为S系列交换机支持三种路径开销计算方法，即 IEEE 802.1d-1998标准方法、IEEE 802.1t 标准方法和华为的私有计算方法。以华为的私有计算方法为例，不同速率的端口路径开销的缺省值不同，具体如表8-5所示。

表8-5 端口所对应的链路速率与端口路径开销缺省值对应表

链路速率	路径开销取值范围	路径开销推荐取值范围	路径开销缺省值
10Mbit/s	1～200 000	200～20 000	2 000
100Mbit/s	1～200 000	20～2 000	200
1Gbit/s	1～200 000	2～200	20
10Gbit/s	1～200 000	2～20	2
10Gbit/s 以上	1～200 000	1～2	1

从上表可以看出，端口速率越高，路径开销值越小。在存在环路的网络环境中，对于链路速率值相对较小的端口，建议将其路径开销值配置相对较大，以使其在生成树算法中被选举成为阻塞端口，阻塞其所在链路，从而可以使速率更高的端口成为指定端口或根端口，以提高网络交换性能。

（5）（可选）配置端口优先级。在参与STP/RSTP生成树计算时，对于处在环路中的交换设备端口，其优先级的高低会影响到是否被选举为指定端口。如果希望将环路中的某交换设备的端口阻塞从而破除环路，则可将其端口优先级值设置比缺省值大（优先级值越大，优先级越小 ），使得在选举过程中成为被阻塞的端口。

（6）启用STP/RSTP功能。在环形网络中一旦启用STP或RSTP，STP、RSTP协议便立即开始生成树计算。而且，诸如交换设备的优先级、端口优先级等参数都会影响到生成树的计算，在计算过程中这些参数的变动可能会导致网络振荡。为了保证生成树计算过程快速而且稳定，必须在交换设备及其端口进行必要的基本配置以后才能启用**STP**或**RSTP**功能 。

（7）（可选）配置端口的收敛方式。当生成树的拓扑结构发生改变时，和它建立映射关系的VLAN的转发路径也将发生变化。此时，交换设备的ARP表中与这些VLAN相关的表项也需要更新。根据对ARP表项的处理方式不同，STP、RSTP的收敛方式分为Fast和Normal两种：在Fast方式下，ARP表将需要更新的表项直接删除；在Normal方式下，ARP表中需要更新的表项快速老化。在normal方式下，交换设备将ARP表中这些表项的剩余存活时间置为0，对这些表项进行老化处理。如果配置的ARP老化探测次数大于零，则ARP对这些表项进行老化探测。

2．具体配置步骤

前面介绍的七大STP和RSTP基本功能主要配置任务的具体配置步骤如表8-6所示。

表8-6 STP/RSTP基本功能配置步骤

配置任务	步骤	命令	说明
	1	**system-view** 例如：<HUAWEI> **system-view**	进入系统视图
配置 STP 或 RSTP 工作模式	2	**stp mode** { **stp** | **rstp** } 例如：[HUAWEI] **stp mode stp**	配置交换机的生成树工作模式。如果选择二选一选项 **stp**，则表示运行 STP 模式；如果选择二选一选项 **rstp**，则表示运行 RSTP 模式 缺省情况下，除 S2700SI 子系列运行模式为 STP 模式外，其他系列为运行 MSTP 模式，MSTP 模式兼容 STP 和 RSTP 模式，可用 **undo stp mode** 命令恢复交换设备的缺省生成树协议工作模式
（可选）配置根桥或备份根桥	3	**stp root** { **primary** | **secondary** } 例如：[HUAWEI] **stp root primary**	配置当前设备为根桥或备份根桥，如果选择二选一选项 **primary**，则配置当前设备为根桥；如果选择二选一选项 **secondary**，则配置当前设备为备份根桥 如果配置为根桥后该设备 BID 中的优先级值自动为 0，并且不能更改；如果配置为备份根桥后该设备 BID 中的优先级值自动为 4 096，且也不能更改 缺省情况下，交换设备不作为任何生成树的根桥或备份根桥，可用 **undo stp root** 命令取消当前交换设备为指定生成树的根桥或备份根桥资格

（续表）

配置任务	步骤	命令	说明
（可选）配置桥优先级	4	**stp priority** *priority* 例如：[HUAWEI] **stp priority** 4096	配置交换设备的桥优先级，取值范围是 0～61 440，步长为 4 096，即**仅可以配置 16 个优先级取值，如 0、4 096、8 192 等，不能随便设**。优先级值越小，则优先级越高，越能成为根桥或备份根桥 缺省情况下，交换设备的桥优先级值为 32 768，可用 **undo stp priority** 命令恢复交换机的桥优先级为缺省值 【注意】如果已经通过执行命令 **stp root primary** 或命令 **stp root secondary** 指定当前设备为根桥或备份根桥，若要改变当前设备的优先级，则需要执行命令 **undo stp root** 去使能根桥或者备份根桥功能，然后执行本命令配置新的优先级数值
（可选）配置端口路径开销	5	**stp pathcost-standard** { **dot1d-1998** | **dot1t** | **legacy** } 例如：[HUAWEI] **stp pathcost-standard dot1d-1998**	配置端口路径开销缺省值的计算方法。命令中的选项说明如下。 • **dot1d-1998**：多选一选项，表示采用 IEEE 802.1D 标准计算方法 • **dot1t**：多选一选项，表示采用 IEEE 802.1t 标准计算方法 • **legacy**：多选一选项，表示采用华为的私有计算方法 缺省情况下，路径开销缺省值的计算方法为 IEEE 802.1t（**dot1t**）标准方法，可用 **undo stp pathcost-standard** 命令恢复路径开销缺省值采用缺省计算方法。且同一网络内所有交换设备的端口路径开销应使用相同的计算方法
	6	**Interface** *interface-type interface-number* 例如：[HUAWEI]**interface** GigabitEthernet **1/0/0**	进入要参与生成树计算的接口视图
	7	**stp cost** *cost* 例如：[HUAWEI-GigabitEthernet1/0/0] **stp cost** 200	设置当前端口的路径开销值，用于桥的根端口选举，值越大，优先级越低。取值范围根据所采用的计算方法的不同而不同。 • 使用华为的私有计算方法时参数 *cost* 的取值范围是 1～200 000 • 使用 IEEE 802.1d 标准方法时参数 *cost* 的取值范围是 1～65 535 • 使用 IEEE 802.1t 标准方法时参数 *cost* 的取值范围是 1～200 000 000 缺省情况下，端口的路径开销值为接口速率对应的路径开销缺省值，可用 **undo stp cost** 命令恢复当前接口的路径开销为缺省值。当采用华为私有计算方法时的缺省值参见表 8-5 【说明】在存在环路的网络环境中，对于链路速率值相对较小的接口，建议将其路径开销值配置相对较大，以使其在生成树算法中被选举成为阻塞端口，阻塞其所在链路，因为开销值越大的接口越将成为阻塞端口
（可选）配置端口优先级	8	**stp port priority** *priority* 例如：[HUAWEI-GigabitEthernet1/0/0] **stp port priority** 64	配置端口的优先级，参与指定端口的选举。参数的 *priority* 取值范围是 0～240，步长为 16，不能随便设置，且优先级值越小，优先级越高，越能成为指定端口 缺省情况下，端口的优先级取值是 128，可用 **undo stp port priority** 命令恢复当前接口的优先级为缺省值
对所有要参与 STP 或者 RSTP 生成树计算的各交换机端口重复以上第 5～第 7 步			

（续表）

配置任务	步骤	命令	说明
启用 STP 或 RSTP	9	**quit** 例如：[HUAWEI-GigabitEthernet1/0/0] **quit**	退出接口视图，返回系统视图
	10	**S2700/3700/5700/6700 系列交换机** （除 **S5710EI** 子系列外）： **bpdu enable** 例如：[HUAWEI] **bpdu enable** **S7700/9300/9700 系列交换机**： **bpdu bridge enable** 例如：[HUAWEI] **bpdu bridge enable**	（可选）使能接口上送 BPDU 报文到 CPU 处理的功能。STP/RSTP 需要通过 BPDU 报文交互来完成生成树计算。因此需要使能接口上送 BPDU 报文到 CPU 处理的功能。但 S5710EI 不支持该命令。在 S5710EI 上，只有使能了 STP 或者 RSTP 的接口才会将相应的 BPDU 报文上送 CPU 处理；否则，接口直接丢弃 BPDU 报文在 **S2700/3700/ 5700/6700 系列交换机中**，缺省情况下，该功能处于使能状态，可用 **bpdu disable** 命令去使能该功能；在 **S7700/9300/9700 系列交换机中**，缺省情况下，接口对收到的 BPDU 报文进行丢弃处理，即去使能该功能，也可用 **bpdu bridge disable** 命令去使能该功能
	11	**stp enable** 例如：[HUAWEI] **stp enable** 或 [HUAWEI-GigabitEthernet1/0/0] **stp enable**	使能交换机的 STP/RSTP 功能。本命令既可在系统视图下全局启用交换机上各端口的STP或者RSTP功能，也可在具体接口视图下仅启用对应接口的 STP 或者 RSTP 功能 缺省情况下，S2700/3700/9300 系列交换机中的 STP/RSTPMSTP 功能处于启用状态，可用 **undo stp enable** 命令去使能交换设备或端口上的 STP/RSTP/MSTP 功能。也可用 **stp disable** 命令去使能交换设备或端口上的 MSTP/RSTP 功能（在 S2700/3700 系列交换机中仅可使用此命令去使能 STP/RSTP/MSTP 功能）；在 S5700/6700/7700/9700 系列交换机中的 MSTP 功能处于禁用状态，可用 **stp enable 或者 undo stp disable** 命令使能交换设备或端口上的 STP/RSTP/MSTP 功能
（可选）配置端口的收敛方式	12	**stp converge** { **fast** | **normal** } 例如：[HUAWEI] **stp converge fast**	配置生成树的收敛方式。命令中的选项说明如下。 • **fast**：二选一选项，指定采用快速方式，ARP 表将需要更新的表项直接删除 • **normal**：二选一选项，指定采用普通模式，仅将 ARP 表中需要更新的表项快速老化 缺省情况下，端口的 STP/RSTP 收敛方式为 normal，可用 **undo stp converge** 命令恢复 STP/RSTP 收敛方式为缺省值。建议选择 normal 收敛方式。若选择 fast 方式，频繁的 ARP 表项删除可能会导致设备 CPU 占用率高达 100%，报文处理超时导致网络振荡

8.4.3 配置影响STP拓扑收敛的参数

虽然说STP不能实现快速收敛，但是诸如网络直径、超时时间、Hello Time定时器、Max Age定时器、Forward Delay定时器等参数会影响其收敛速度。本节要具体介绍这些参数的具体配置方法，在配置影响STP拓扑收敛的参数之前，需要完成上节介绍的STP 基本功能配置。下面先具体了解这些参数的作用。

1．影响STP拓扑收敛的参数

（1）STP网络直径。交换网络中任意两台终端设备都通过特定路径彼此相连，这些路径由一系列的交换设备构成。网络直径就是指交换网络中任意两台终端设备间的最大交换设备数。网络直径越大，说明网络的规模越大。但是这里的网络直径也不是随便设的，因为如果网络直径设置不合理，可能会引起网络收敛速度慢，影响用户的正常通信。根据当前的网络规模，设置合适的网络直径（通常不要超过7个设备），可以帮助加快网络收敛速度。建议同一环网中的所有交换设备配置相同的网络直径。

（2）STP超时时间。在运行STP生成树算法的交换网络中，如果交换设备在配置的超时时间内没有收到上游设备发送的BPDU就认为此上游设备已经出现故障，本设备会重新进行生成树计算。可能由于上游设备繁忙，有时设备在较长的时间内收不到该上游设备发送的 BPDU。在这种情况下一般不应该重新进行生成树计算，因此在稳定的网络中，可以配置超时时间，以减少网络资源的浪费。

（3）STP定时器。在 STP生成树的计算过程中，用到了 Forward Delay、Hello Time和Max Age 3个定时

器参数，具体参见本章 8.1.3节介绍。在配置这 3个定时器参数时，同一环网中的设备建议配置一致的定时器值。但是，通常情况下，不建议通过本配置直接调整上述3个时间参数，而是建议通过调整网络直径，使生成树协议自动调整这3个定时器参数的值。当网络直径取缺省值时，这 3 个定时器参数也分别取其各自的缺省值。

（4）影响链路聚合带宽最大连接数。接口的路径开销是生成树计算的重要依据，路径开销值改变时，会重新进行生成树计算。而接口的路径开销是受带宽影响的，因此可以通过改变接口带宽来影响生成树的计算。当接口是 Eth-Trunk 的聚合接口时，可配置链路聚合带宽最大连接数，以选择适当的聚合链路。当然，这里配置的影响带宽的最大连接数仅影响生成树协议计算接口的链路开销，并不影响实际链路带宽。Eth-Trunk接口在转发流量时的实际带宽仍然是由活动接口数决定的。

如图8-22所示，设备A与设备B通过两条Eth-Trunk链路相连，Eth-Trunk1含有3条状态为Up的成员链路，Eth–Trunk2含有2条状态为Up的成员链路。假设每条成员链路的带宽都相同，且设备A被选举为根桥，因为Eth-Trunk1的带宽大于Eth-Trunk2的带宽。STP计算后，设备B上 Eth-Trunk1端口被选为Root port，Eth-Trunk2端口被选为Alternate port。

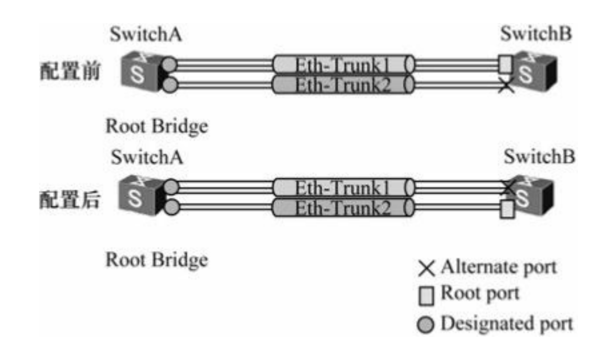


图8-22 影响链路聚合带宽的最大连接数示例

但当配置Eth-Trunk1接口影响带宽的最大连接数为1后，STP计算的Eth-Trunk1接口的路径开销大于Eth-Trunk2的开销，将会重新进行生成树计算，设备B上Eth-Trunk1接口将变为Alternate port，Eth-Trunk2变为Root port。

2．具体配置步骤

下面针对以上4项影响STP拓扑收敛的参数的具体配置步骤进行介绍，如表8-7所示。注意，这里的参数配置没有严格的先后顺序。

表8-7 STP参数配置步骤

配置任务	步骤	命令	说明
配置 STP 网络直径	1	**system-view** 例如：<HUAWEI> **system-view**	进入系统视图
	2	**stp bridge-diameter** *diameter* 例如：[HUAWEI] **stp bridge-diameter** 5	配置网络直径（指任意两个交换设备之间的交换设备个数的最大值），取值范围为 2～7 的整数 执行本命令后，交换设备会自动根据配置的网络直径设置 Hello Time、Forward Delay 与 Max Age 3 个时间参数为较优值，且在配置文件中会出现 Forward Delay 与 Max Age 两个时间参数的具体配置值，所以建议通过本命令配置的网络直径自动去配置 Forward Delay 时间、Hello Time 时间以及 Max Age 时间，因为交换设备会自动根据网络直径计算出 Forward Delay 时间、Hello Time 时间以及 Max Age 时间的最优值 缺省情况下，生成树的网络直径为 7，可用 **undo stp bridge-diameter** 命令恢复网络直径为缺省值
配置 STP 超时时间	3	**stp timer-factor** *timer-factor* 例如：[HUAWEI] **stp timer-factor** 5	配置未收到上游的 BPDU 就重新开始生成树计算的超时时间的时间因子（或者 Hello Time 的时间倍数），取值范围为 1～10 的整数。参数 *factor* 的值配置越小，表示交换设备重新进行生成树拓扑计算的超时时间越短，则错误判断上游交换设备已经出现故障的概率越大；参数 *factor* 的值配置越大，表示交换设备重新进行生成树拓扑计算的超时时间越长，会导致在上游交换设备已经出现故障的情况下，接口中断流量的概率越大 在 S2700/3700/9300 系列交换机中，超时时间＝Hello Time × Timer Factor，在 S5700/6700/7700/9700 系列交换机中，超时时间＝Hello Time × 3 × Timer Factor 缺省情况下，Timer Factor 的取值为 3，但因为不同系列交换机的超时时间计算公式不一样，所以最终的缺省超时时间值也不一样：在 S5700/6700/7700/9700 系列交换机中，设备未收到上游的 BPDU 就重新开始生成树计算的缺省超时时间是 Hello Timer 的 3 倍，而在 S5700/6700/7700/9700 系列交换机中，设备未收到上游的 BPDU 就重新开始生成树计算的缺省超时时间是 Hello Timer 的 9 倍

（续表）

配置任务	步骤	命令	说明	
配置 STP 定时器	4	**stp timer forward-delay** *forward-delay* 例如：[HUAWEI] **stp timer forward-delay** 2000	配置设备的 Forward Delay 时间，取值范围为（400～3000）的整数厘秒，步长是 100 在 STP 协议中，接口由 Discarding 状态转向 Forwarding 状态时要经历两个 Forward Delay 时间的延迟。在根桥上配置的 Forward Delay 时间将作为整个生成树内所有桥的 Forward Delay 时间 缺省情况下，设备的 Forward Delay 时间是 1 500 厘秒，可用 **undo stp timer forward-delay** 恢复 Forward Delay 时间为缺省值	根桥的 Hello Time、Forward Delay 和 Max Age 3 个定时器参数取值之间应该满足如下公式，否则网络会频繁振荡： • 2 × (Forward Delay − 1.0 second) ≥Max Age • Max Age ≥ 2× (Hello Time + 1.0 second) 建议使用前面介绍的 **stp bridge-diameter** 命令配置网络直径，交换设备会自动根据网络直径计算出 Hello Time、Forward Delay 以及 Max Age 3 个定时器参数的较优值
	5	**stp timer hello** *hello-time* 例如：[HUAWEI] **stp timer hello** 200	配置设备的 Hello Time 时间，取值范围为（100～1 000）的整数厘秒，步长为 100 在运行 STP 算法的网络中，以 Hello Time 为周期，交换设备会定时向处于同一棵生成树的其他设备发送 BPDU，以此来维护生成树的稳定。通过执行本命令设置 BPDU 发送间隔，维护网络拓扑结构的稳定。如果交换设备在超时时间内没有收到上游交换设备发送的 BPDU，则生成树会重新进行计算 **在根桥上配置的定时器 Hello Timer 的时间将通过 BPDU 传递下去，所以会成为整棵生成树内所有交换设备的定时器 Hello Timer 的时间** 缺省情况下，设备的 Hello Time 时间是 200 厘秒，可用 **undo stp timer hello** 命令恢复交换设备发送 BPDU 的时间间隔为缺省值	
	6	**stp timer max-age** *max-age* 例如：[HUAWEI] **stp timer max-age** 1000	配置设备的 Max Age 时间，取值范围为（600～4000）的整数厘秒，步长为 100 在运行 STP 算法的网络中，交换设备会根据端口的 Max Age 时间判断从上游交换设备收到的 BPDU 是否超时。如果 BPDU 超时，交换设备将该 BPDU 老化，同时阻塞接收该 BPDU 的端口，并发出以自己为根桥的 BPDU。这种老化机制可以有效控制生成树的半径。通过执行本命令设置 Max Age 定时器时间的大小控制存储 BPDU 的超时时间 缺省情况下，设备的 Max Age 时间是 2 000 厘秒，可用 **undo stp timer max-age** 命令恢复交换设备端口的 BPDU 老化时间为缺省值	

（续表）

配置任务	步骤	命令	说明
配置影响生成树计算的链路聚合带宽最大连接数	7	**interface eth-trunk** *trunk-id* 例如：[HUAWEI] **interface eth-trunk** 1	进入 Eth-Trunk 接口视图
	8	**max bandwidth-affected-linknumber** *link-number* [HUAWEI-Eth-Trunk1] **max bandwidth-affected-linknumber** 5	配置影响链路聚合带宽接口数目的上限阈值，除 S2700SI 子系列交换机的取值范围为 1～4 的整数外，其他支持 VRP 系统的 S 系列交换机的取值范围为 1～8 的整数 缺省情况下，影响链路聚合带宽的最大连接数除 S2700SI 子系列交换机为 4 外，其他均为 8
查看配置结果	9	**display stp** [**interface** *interface-type interface-number* | **slot** *slot-id*] [**brief**]	查看生成树的状态信息与统计信息。命令中的参数说明如下。 ● *interface-type interface-number*：二选一可选参数，指定要查看生成树状态信息和统计信息的接口 ● *slot-id*：二选一可选参数，指定要查看生成树状态信息和统计信息的槽位号 ● **brief**：可选项，指定仅显示生成树的状态和统计信息摘要，如果不选择此可选项，则查看生成树的状态和统计的详细信息 如果不指定以上两个可选参数，则查看交换机上所有端口的生成树的状态和统计信息

【示例 】当没有在交换机上执行 stp enable命令时，通过display stp命令在交换机上显示的生成树的状态和统计信息如下所示，输出信息字段说明如表8-8所示。

<HUAWEI >display stp

Protocol Status :Disabled

Protocol Standard :IEEE 802.1s

Version :3

CIST Bridge Priority :32768

MAC address :00e0-6343-6800

Max age(s) :20

Forward delay(s) :15

Hello time(s) :2

Max hops :20

Share region-configuration :Enabled

表8-8 在未启用STP/RSTP功能时执行display stp命令的输出信息字段说明

字段	说明
Protocol Status	显示 STP/RSTP 协议状态：Disabled 为去使能，Enabled 为使能
Protocol Standard	显示设备当前运行的生成树协议标准
Version	显示设备当前运行的生成树协议版本：0 代表 STP，2 代表 RSTP，3 代表 MSTP
CIST Bridge Priority	显示交换设备在 CIST（公共内部生成树）中的优先级，仅运行 MSTP 协议时显示
MAC address	显示交换设备的 MAC 地址
Max age(s)	显示 BPDU 最大生存时间
Forward delay(s)	显示端口状态迁移的延时
Hello time(s)	显示根交换设备发送 BPDU 的周期
Max hops	显示 MST 域中的最大跳数，仅当 MSTP 协议时显示
Share region-configuration	显示共享 MST 区域配置，enabled 表示共享进程 0 的配置

8.4.4 STP配置示例

STP方面的配置就是前面两节介绍的那些，比较简单，大多数情况下是不需要什么配置的，因为STP功能缺省是启用的，只不过在需要使用STP协议时，配置生成树模式为STP，至于其他的均可根据需要选择配置，包括根桥、备份根桥的指定。当然，通常是建议手工指定根桥和备份根桥，这样可以使自己更加清楚

自己交换网络的拓扑结构。

本示例拓扑结构如图8-23所示，当前网络中存在由SwitchA、SwitchB、SwitchC和 SwitchD 构成的环路，因为在 SwitchA与SwitchD之间，以及SwitchB与SwitchC之间都存在冗余链路（本来这些链路都是可以不要的）现在这些交换机上都运行STP 协议，通过彼此交互信息发现网络中的环路，并有选择地对某个端口进行阻塞，最终将环形网络结构修剪成无环路的树形网络结构，从而防止报文在环形网络中不断循环，避免设备由于重复接收相同的报文造成处理能力下降。

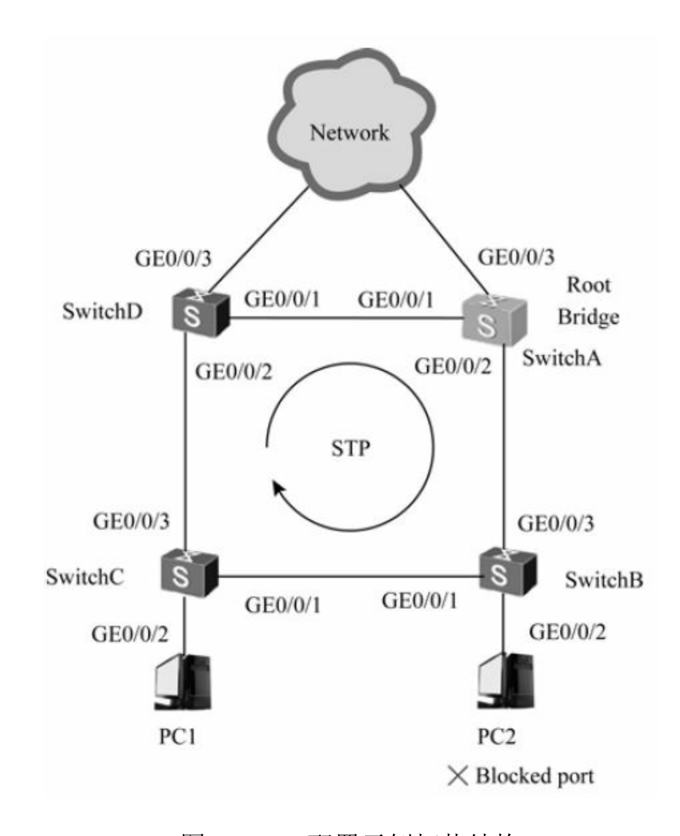


图8-23 STP配置示例拓扑结构

1．配置思路分析

本示例在参数方面没有特别的要求，所以实际本示例仅需要针对8.4.2节的介绍配置 STP 的基本功能即可。基本配置思路如下（仅针对环网结构中的4台交换机）。

（1）配置环网中的 4 台交换机的生成树协议工作在STP模式。

（2）配置根桥和备份根桥设备，此处可以指定SwitchA为根桥，SwitchD为备份根桥。

（3）配置端口的路径开销值，实现将该端口阻塞。此处可以加大SwitchC的GE0/0/1端口的开销值，以阻塞该端口，使得数据不能从该端口发送。

（4）在四台交换机上使能STP功能。但与PC机相连的端口不用参与STP计算，建议将其去使能STP。

2．具体配置步骤

下面具体介绍以上配置任务中的具体配置步骤。注意，要在对应交换机上配置。

（1）在4台环网结构中的交换机上配置STP工作模式。因为4台交换机上的配置方法完全一样，故下面仅以SwitchA交换机上的配置为例进行介绍，其他交换机的配置参见即可。

```
<HUAWEI>system-view
```

[HUAWEI] sysname SwitchA

[SwitchA] stp mode stp

（2）配置SwitchA为根桥，SwitchD为备份根桥。

[SwitchA] stp root primary

[SwitchD] stp root secondary

（3）配置端口的路径开销计算方法，同时将SwitchC上的GE0/0/1端口的开销值增大（大于对应类型端口的路径开销缺省值），实现将该端口的阻塞。

端口路径开销值取值范围由路径开销计算方法决定，这里以使用华为私有计算方法为例。同样因为4台交换机上的路径开销计算方法的配置方法完全一样，在此仅以SwitchA上的配置为例进行介绍。但同一网络内所有交换设备的端口路径开销应使用相同的计算方法。

[SwitchA] stp pathcost-standard legacy

然后增大SwitchC上的GE0/0/1端口的开销值，此处为20 000（千兆端口的缺省值为2）。

[SwitchC] interfacegigabitethernet 0/0/1

[SwitchC-GigabitEthernet0/0/1] stp cost 20000

（4）在 4 台交换机使能 STP 功能，以消除二层环路。但在此之前要先去使能连接PC上的端口（如SwitchB的GE0/0/2端口和SwitchC的GE0/0/2端口）上的STP功能。

[SwitchB] interfacegigabitethernet 0/0/2

[SwitchB-GigabitEthernet0/0/2] stp disable

[SwitchB-GigabitEthernet0/0/2] quit

[SwitchC] interfacegigabitethernet 0/0/2

[SwitchC-GigabitEthernet0/0/2] stp disable

[SwitchC-GigabitEthernet0/0/2] quit

然后在4台交换机上全局使能STP。同样因为四台交换机上的使能方法的配置方法完全一样，在此仅以SwitchA上的配置为例进行介绍。

[SwitchA] stp enable

以上配置完成后，过段时间，在网络计算稳定后执行以下命令，以验证配置结果。

在SwitchA上执行display stp brief命令查看端口状态和端口的保护类型，结果如下。

[SwitchA] display stp brief

MSTID	Port	Role	STP State	Protection
0	GigabitEthernet0/0/1	DESI	FORWARDING	NONE
0	GigabitEthernet0/0/2	DESI	FORWARDING	NONE

将SwitchA配置为根桥后，与SwitchB、SwitchD相连的GE0/0/2和GE0/0/1端口在生成树计算中被选举为指定端口。可通过在 SwitchB 上执行 display stp interface gigabitethernet 0/0/1 brief命令查看端口GigabitEthernet0/0/1状态来验证，结果如下，从中可以看出GE0/0/1端口在生成树选举中已成为指定端口，处于Forwarding状态。

[SwitchB] display stp interfacegigabitethernet0/0/1 brief

MSTID	Port	Role	STP State	Protection
0	GigabitEthernet0/0/1	DESI	FORWARDING	NONE

同样可在在SwitchC上执行display stp brief命令查看端口状态，结果如下，从中可以看出GE0/0/3端口在生成树选举中成为根端口，处于Forwarding状态，而GE0/0/1端口在生成树选举中成为Alternate端口，处于

Discarding状态。

[SwitchC] display stp brief

MSTID	Port	Role	STP State	Protection
0	GigabitEthernet0/0/1	ALTE	DISCARDING	NONE
0	GigabitEthernet0/0/3	ROOT	FORWARDING	NONE

通过以上查看操作就可以验证以上配置是正确、成功的。

8.4.5 配置影响RSTP拓扑收敛的参数

RSTP在STP基础上进行改进之后，通过配置端口的链路类型、端口是否支持快速迁移机制等，实现快速收敛。其本身的基本功能配置仍如STP基本功能配置方法差不多，已在8.4.2节进行了介绍。在进行本节配置之前，需完成RSTP基本功能配置。

1．影响拓扑收敛的参数

在RSTP中，影响拓扑收敛的参数除了STP中介绍的网络直径、超时时间、三个定时器、影响生成树计算的链路聚合带宽最大连接数这4个外，还有端口的链路类型、端口的最大发送速率、是否执行MCheck操作、边缘端口和BPDU报文过滤功能启用等几个方面。下面介绍这几个在RSTP新增的影响拓扑收敛的参数。

（1）端口的链路类型

点对点链路可帮助实现快速收敛。在RSTP中，如果与点对点链路相连的两个端口为根端口或者指定端口，则端口可以通过传送同步报文（Proposal 报文和 Agreement 报文）快速迁移到转发状态，减少了不必要的转发延迟时间。

（2）端口的BPDU报文最大发送速率

接口在每个Hello Time时间内BPDU的最大发送数目值越大，表示单位时间内发送的BPDU越多，占用的系统资源也越多。适当地配置该值可以限制接口发送BPDU的速度，防止在网络拓扑动荡时，RSTP占用过多的带宽资源。

（3）执行MCheck操作

在运行RSTP的设备上，如果某个接口和另一台运行STP的设备连接，则该接口会自动迁移到STP兼容工作模式。但如果某一时间运行STP的设备被关机或移走（还可能是因为原来STP的交换设备切换为RSTP模式），原来自动迁移到STP兼容工作模式的接口无法自动迁移回RSTP模式。这时就需要在该接口上执行MCheck操作，将接口手动迁移到RSTP模式。

（4）边缘端口和BPDU报文过滤功能

在RSTP里面，位于整个网络的边缘（即不再与其他交换设备连接，而是直接与终端设备直连）的端口叫做边缘端口。边缘端口不接收处理配置BPDU报文，不参与RSTP运算，可以由Disable直接转到Forwarding状态，且不经历时延，就像在端口上将RSTP禁用。

配置为边缘端口后，端口仍然会发送BPDU报文，这可能导致BPDU报文发送到其他网络，引起其他网络产生振荡。因此可以配置边缘端口的BPDU报文过滤功能，使边缘端口不处理、不发送BPDU报文。

边缘端口和BPDU报文过滤功能可以在系统视图下全局配置，也可在具体端口的接口视图下配置，通常是在具体端口的接口视图下，因为不可能交换机上所有端口都是边缘端口。当然如果交换机上大多数端口为边缘端口，则可先通过全局配置使所有端口成为边缘端口，然后对不要配置为边缘的少数端口恢复为非边缘端口类型即可。

全局配置后，设备上所有的端口不会主动发送BPDU报文，且均不会主动与对端设备直连端口协商，所

有端口均处于转发状态；在接口配置后，对应端口将不处理、不发送BPDU报文，无法成功与对端设备直连端口协商STP协议状态。

2．具体配置步骤

因为RSTP中前面四项参数与STP中的对应参数配置方法完全一样，所以可直 接参见8.4.3节表8-7即可，下面具体介绍前面后面四项影响RSTP拓扑收敛的各项参数的具体配置步骤，如表 8-9 所示，但要注意，这些参数的配置也是根据需要可选的。

表8-9 RSTP参数配置步骤

<table>
<tr><th>配置任务</th><th>步骤</th><th>命令</th><th colspan="3">说明</th></tr>
<tr><td rowspan="3">配置端口的链路类型</td><td>1</td><td>system-view
例如：<HUAWEI> system-view</td><td colspan="3">进入系统视图</td></tr>
<tr><td>2</td><td>interface interface-type interface-number
例如：[HUAWEI]interface gigabitethernet 0/0/1</td><td colspan="3">进入参与生成树协议计算的接口视图。此接口为指定端口</td></tr>
<tr><td>3</td><td>stp point-to-point { auto | force-false | force-true }
例如：[HUAWEI-GigabitEthernet0/0/1]stp point-to-point force-true</td><td colspan="3">配置指定端口的链路类型，命令中的参数说明如下。
• auto：多选一选项，指定由生成树协议自动检测与该端口相连的链路是否是点到点链路
• force-false：多选一选项，指定与当前端口相连的链路不是点到点链路
• force-true：多选一选项，指定与当前端口相连的链路是点到点链路
如果当前以太网端口工作在全双工模式，则当前端口相连的链路是点到点链路，选择 force-true 选项，以实现快速收敛。如果当前以太网端口工作在半双工模式，可通过执行选择 force-true 选项强制链路类型为点对点链路，实现快速收敛
缺省情况下，指定端口自动识别是否与点对点链路相连，可用 undo stp point-to-point 命令恢复指定端口的链路类型为缺省类型</td></tr>
<tr><td>配置端口的 BPDU 报文最大发送速率</td><td>4</td><td>stp transmit-limit packet-number
例如：[HUAWEI-GigabitEthernet0/0/1] stp transmit-limit 5
或
例如：[HUAWEI]stp transmit-limit 5</td><td colspan="3">配置当前端口（在接口视图下配置）或本设备上所有端口（在系统视图下配置）在单位时间内 BPDU 的最大发送数目，取值范围为 1～255 的整数
缺省情况下，端口每秒 BPDU 的最大发送数目为 6，可用 undo stp transmit-limit 命令恢复当前端口（在接口视图下配置时）或本设备上所有端口（在系统视图下配置时）每秒发送 BPDU 的最大数目为缺省值</td></tr>
<tr><td rowspan="3">配置设备执行 MCheck 操作</td><td rowspan="2">5</td><td>stp mcheck
例如：[HUAWEI-GigabitEthernet0/0/1] stp mcheck</td><td>在具体接口视图下执行 MCheck 操作，将当前端口执行自动迁移回原来的 RSTP 模式</td><td>接口配置方式</td><td rowspan="3">二选一</td></tr>
<tr><td>system-view
例如：[HUAWEI] system-view</td><td>进入系统视图</td><td rowspan="2">全局配置方式</td></tr>
<tr><td>6</td><td>stp mcheck
例如：[HUAWEI] stp mcheck</td><td>全局执行 MCheck 操作，对交换设备上所有端口执行自动迁移回原来的 RSTP 模式</td></tr>
</table>

（续表）

配置任务	步骤	命令	说明		
配置边缘端口和BPDU报文过滤功能	7	**stp edged-port default** 例如：[HUAWEI]**stp edged-port default**	配置当前设备上所有端口为边缘端口。端口配置成边缘端口后，如果收到 BPDU 报文，交换设备会自动将边缘端口设置为非边缘端口，并重新进行生成树计算。为防止攻击者仿造 BPDU 报文导致边缘端口属性变成非边缘端口，建议在系统视图下执行 **stp bpdu-protection** 命令配置交换设备的 BPDU 保护功能。配置 BPDU 保护功能后，如果边缘端口收到 BPDU 报文，边缘端口将会被 shutdown，边缘端口属性不变 缺省情况下，设备的所有端口为非边缘端口，可用 **undo stp edged-port default** 命令恢复交换设备所有端口为非边缘端口	全局配置方式	二选一
	8	**stp bpdu-filter default** 例如：[HUAWEI] **stp bpdu-filter default**	配置当前设备上所有端口为 BPDU filter 端口。在接口视图下使用命令 **stp bpdu-filter disable** 命令将不需要配置成 BPDU filter 端口的端口恢复为非 BPDU filter 端口 缺省情况下，设备的所有端口为非 BPDU filter 端口，可用 **undo stp bpdu-filter default** 命令配置当前设备上所有端口为非 BPDU filter 端口		
	7	**Interface** *interface-type interface-number* 例如：[HUAWEI] **interface** gigabitethernet 0/0/1	进入参与生成树协议计算的以太接口视图	接口配置方式	
	8	**stp edged-port enable** 例如：[HUAWEI-GigabitEthernet0/0/1] **stp edged-port enable**	将端口配置成边缘端口。与终端相连的端口不用参与生成树计算，可以通过执行本命令将当前端口配置成边缘端口，该端口便不再参与生成树计算，从而加快网络拓扑的收敛时间，加强网络的稳定性 但是当通过本命令将当前端口配置成边缘端口后，仍然会发送 BPDU 报文，这可能导致 BPDU 报文发送到其他网络，引起其他网络产生振荡。这里还要通过下一步的 **stp bpdu-filter enable** 命令解决 缺省情况下，交换设备的所有端口都是非边缘端口，可用 **stp edged-port disable** 或 **undo stp edged-port** 命令配置当前端口为缺省的非边缘端口属性		

（续表）

配置任务	步骤	命令	说明		
配置边缘端口和BPDU报文过滤功能	9	**stp bpdu-filter enable** 例如：[HUAWEI-GigabitEthernet0/0/1] **stp bpdu-filter enable**	配置当前端口为 BPDU filter 端口。配置本命令后，该端口将无法成功与对端设备直连端口协商 STP 协议状态 缺省情况下，设备的所有端口为非 BPDU filter 端口，可用 **stp bpdu-filter disable** 或 **undo stp bpdu-filter** 命令配置当前端口为非 BPDU filter 端口	接口配置方式	二选一

8.4.6 配置RSTP保护功能

华为公司的数据通信设备支持如表8-3所示的RSTP保护功能，用户可根据实际环境任选其中一个或多个保护功能配置。当然，也可以不配置这些保护功能。

RSTP保护功能配置步骤如表8-10所示。各保护功能的配置没有严格的先后顺序。

表8-10 RSTP保护功能的配置步骤

配置任务	步骤	命令	说明
配置 BPDU 保护功能	1	**system-view** 例如：<HUAWEI> **system-view**	进入系统视图
	2	**stp bpdu-protection** 例如：[HUAWEI]**stp bpdu-protection**	配置边缘端口的 BPDU 保护功能。配置 BPDU 保护功能后，如果边缘端口收到 BPDU 报文，边缘端口将会被 error-down，边缘端口属性不变 如果用户希望被 error-down 的边缘端口可自动恢复，可通过在系统视图下执行 **error-down auto-recoverycause bpdu-protection interval** *interval-value* 命令，配置使能端口自动恢复功能，并设置延迟时间，使被关闭的端口经过延时时间后能够自动恢复。但在配置自动恢复功能时需要注意以下几个方面。 • 缺省情况下，未使能处于 error-down 状态的接口状态自动恢复为 Up 的功能，所以没有缺省延迟时间值。当用户配置本命令时，必须指定恢复延迟时间 • 参数 *interval-value* 取值范围为（30～86 400）的整数秒，取值越小表示接口的管理状态自动恢复为 Up 的延迟时间越短，接口 Up/Down 状态振荡频率越高；取值越大表示接口的管理状态自动恢复为 Up 的延迟时间越长，接口流量中断时间越长 • 自动恢复功能仅对配置了本命令之后发生 error-down 的端口有效，对配置此命令之前已经 error-down 的接口不生效 缺省情况下，设备的 BPDU 保护功能处于去使能状态，可使用 **undo stp bpdu-protection** 命令去使能设备的 BPDU 保护功能

（续表）

配置任务	步骤	命令	说明
配置 TC 保护功能	3	**stp tc-protection** 例如：[HUAWEI] **stp tc-protection**	使能交换设备对 TC 类型 BPDU 报文的保护功能。执行本命令，在单位时间内交换设备处理 TC 类型的 BPDU 报文的次数可通过下一步的 **stp tc-protection threshold** 命令配置 缺省情况下，交换设备的 TC 保护功能处于关闭状态，可用 **undo stp tc-protection** 命令去使能设备对 TC 类型 BPDU 报文的保护功能
	4	**stp tc-protection threshold** *threshold* 例如：[HUAWEI] **stp tc-protection threshold** 10	配置交换设备在收到 TC 类型 BPDU 报文后，单位时间内处理 TC 类型 BPDU 报文，并立即刷新转发表项的阈值，参数 *threshold* 的取值范围为 1～255 的整数 【说明】如果在单位时间内，交换设备收到拓扑变化报文数量大于配置的阈值，那么设备只会处理阈值指定的次数。对于其他超出阈值的拓扑变化报文，定时器到期后设备只对其处理一次。这样可以避免频繁地删除 MAC 地址表项和 ARP 表项，从而达到保护设备的目的 缺省情况下，单位时间内处理 TC 类型 BPDU 报文并立即刷新转发表项的缺省值是 1，可用 **undo stp tc-protection threshold** 命令恢复缺省值
配置端口的 Root 保护功能	5	**interface** *interface-type interface-number* 例如：[HUAWEI]**interface** gigabitethernet 0/0/1	进入指定端口的进口视图
	6	**stp root-protection** 例如：[HUAWEI-GigabitEthernet0/0/1] **stp root-protection**	配置以上指定端口（**只能在指定端口下配置**）的 Root 保护功能 【说明】在指定端口使能根保护功能后，收到优先级更高的 BPDU 时该端口状态将进入 Discarding 状态，不再转发报文。在经过一段时间（通常为两倍的 Forward Delay）后，如果端口一直没有再收到优先级较高的 BPDU，该指定端口会自动恢复到正常的 Forwarding 状态，但**配置了根保护的端口不可以再配置下面将要介绍的环路保护功能** 缺省情况下，端口的 Root 保护功能处于去使能状态，可用 **undo stp root-protection** 命令去使能当前指定端口的根保护功能
配置端口的环路保护功能	7	**quit** 例如：[HUAWEI-GigabitEthernet0/0/1] **quit**	退出以上指定端口的接口视图，返回系统视图
	8	**Interface** *interface-type interface-number* 例如：[HUAWEI]**interface** gigabitethernet 0/0/2	进入根端口或者 Alternate 端口的接口视图

（续表）

配置任务	步骤	命令	说明
配置端口的环路保护功能	9	**stp loop-protection** 例如：[HUAWEI-GigabitEthernet0/0/2]**stp loop-protection**	配置交换设备根端口或 Alternate 端口的环路保护功能，**不能在指定端口下配置** 【说明】在启动了环路保护功能后，如果根端口或 Alternate 端口长时间收不到来自上游设备的 BPDU 报文时，则向网管发出通知信息（此时根端口会进入 Discarding 状态），而阻塞端口则会一直保持在阻塞状态，不转发报文，从而不会在网络中形成环路。直到根端口或 Alternate 端口收到 BPDU 报文，端口状态才恢复正常为 Forwarding 状态。配置本命令后就可以防止这种现象发生 由于 Alternate 端口是根端口的备份端口，如果交换设备上有 Alternate 端口，需要在根端口和 Alternate 端口上同时配置环路保护。但配置了根保护的端口不可以再配置环路保护功能 缺省情况下，端口的环路保护功能处于关闭状态，可使用 **undo stp loop-protection** 命令去使能当前端口的环路保护功能

8.4.7 配置设备支持和其他厂商设备互通的参数

在RSTP协议中，网络收敛主要依靠P/A协商机制，但不同厂商设备所支持的P/A机制工作方式不完全一样。为了实现华为公司的数据通信设备与其他厂商设备互通，需要根据其他厂商设备支持的P/A机制选择端口的快速迁移方式。目前，华为S系列交换机的RSTP P/A机制支持以下两种模式。

1．普通方式（Normal mode）

这是一种正常的P/A机制工作方式，双方是通过一对Proposal/Agreement报文进行协商，收到 Proposal 报文的端口为根端口，并自动进入到 Forwarding 状态，而收到Agreement报文的端口为指定端口，也自动进入Forwarding状态。具体流程如下。

（1）上游设备发送Proposal报文，请求进行快速迁移，下游设备在接收后把与上游设备相连的端口设置为根端口，并阻塞所有非边缘端口，然后根端口自动进入Forwarding状态。

（2）然后下游设备回应Agreement报文，上游设备在接收后把与下游设备相连的端口设置为指定端口，指定端口进入Forwarding状态。

2．增强模式（Enhanced mode）

这种方式特别适用于不同厂商设备之间的P/A协商。在这种工作方式中，上游设备发送的 Proposal 报文，在到达下游非同一厂商的设备的根端口时可能不能马上进入Forwarding状态，这时上游设备再发送一个Agreement报文，强制下游设备的根端口进入Forwarding状态。这时下游设备的根端口才可以发送Agreement报文，响应上游设备发送的Proposal报文，使上游设备的指定端口也进入Forwarding状态。具体流程如下。

（1）首先上游设备发送Proposal报文（Flages字段Bit1位置1的BPDU报文），请求进行快速迁移，下游设备在接收后把与上游设备相连的端口设置为根端口，并阻塞所 有非边缘端口（包括根端口）。

（2）然后上游设备继续发送Agreement报文（Flages字段Bit6位置1的BPDU报文），下游设备在接收后强制根端口转为Forwarding状态。

（3）最后下游设备回应Agreement报文，上游设备在接收后把与下游设备相连的端口设置为指定端口，并进入Forwarding状态。

在运行生成树的通信网络中，如果华为公司的数据通信设备与其他厂商设备混合组网，可能会因为与其他厂商设备的 Proposal/Agreement 机制不同导致互通失败。需要根据其他厂商设备的 Proposal/Agreement 机制，选择接口使用增强的快速迁移机制还是普通的快速迁移机制。

配置的方法很简单，只需要直接在其他厂商设备的端口 的接口视图下执行 **stp no-agreement-check** 命令配置端口使用普通的快速迁移方式。缺省情况下，端口使用增强的快速迁移机制，可用**undo stp no-agreement-check** 命令配置当前接口使用增强的快速迁移机制。

8.4.8 RSTP功能配置示例

本示例仍以8.4.4节的图8-23为例进行介绍，环网中的SwitchA、SwitchB、SwitchC和 SwitchD 4台交换机都运行RSTP协议，通过彼此交互信息发现网络中的环路，并有选择地对某个端口进行阻塞，最终将环形网络结构修剪成无环路的树形网络结构，从而防止报文在环形网络中不断循环，避免设备由于重复接收相同的报文造成处理能力下降。

1．配置思路分析

本示例的配置方法与 8.4.4.节 STP 配置示例的配置方法是差不多的，只是在 RSTP中可以把连接PC的端口直接配置为边缘端口，过滤BPDU报文，同时可配置一些保护功能，如本示例就可以使用RSTP根保护功能，使得SwitchA总是为根桥。具体的配置思路如下（仅针对环网结构中的4台交换机）。

（1）配置环网中的4台交换机的生成树协议工作在RSTP模式。

（2）配置根桥和备份根桥设备，此处可以指定SwitchA为根桥，SwitchD为备份根桥。

（3）配置端口的路径开销值，实现将该端口阻塞。此处可以加大SwitchC的GE0/0/1端口的开销值，以阻塞该端口，使得数据不能从该端口发送。

（4）在4台交换机上使能RSTP功能。但与PC机相连的端口不用参与RSTP计算，配置为边缘端口，并配置BPDU过滤。

（5）在SwitchA的GE10/1和GE1/0/2端口上启用根保护功能，使它们总为指定端口，从而使SwitchA总为根桥。

2．具体配置步骤

下面具体介绍以上配置任务中的具体配置步骤。注意，要在对应交换机上配置。

（1）在4台环网结构中的交换机上配置STP工作模式。因为4台交换机上的配置方法完全一样，故下面仅以SwitchA交换机上的配置为例进行介绍，其他交换机的配置参见即可。

```
<HUAWEI>system-view
[HUAWEI] sysname SwitchA
[SwitchA] stp mode rstp
```

（2）配置SwitchA为根桥，SwitchD为备份根桥。

```
[SwitchA] stp root primary
[SwitchD] stp root secondary
```

（3）配置端口的路径开销计算方法，同时将SwitchC上的GE0/0/1端口的开销值增大（大于对应类型端口的路径开销缺省值），实现将该端口阻塞。

端口路径开销值取值范围由路径开销计算方法决定，这里以使用华为私有计算方法为例。同样因为 4 台交换机上的路径开销计算方法的配置方法完全一样，在此仅以SwitchA 上的配置为例进行介绍。但同一网络内所有交换设备的端口路径开销应使用相同的计算方法。

```
[SwitchA] stp pathcost-standard legacy
```

然后增大 SwitchC上的GE0/0/1端口的开销值，此处为 20 000（千兆端口的缺省值为2）。

```
[SwitchC] interfacegigabitethernet 0/0/1
[SwitchC-GigabitEthernet0/0/1] stp cost 20000
```

（4）把连接PC上的端口（如SwitchB的GE0/0/2端口和SwitchC的GE0/0/2端口）配置为边缘端口并配置BPDU过滤。然后在4台交换机上使能RSTP功能，以消除二层环路。

```
[SwitchB] interface gigabitethernet 0/0/2
[SwitchB-GigabitEthernet0/0/2] stp edged-port enable
[SwitchB-GigabitEthernet0/0/2] stp bpdu-filter enable
[SwitchB-GigabitEthernet0/0/2] quit
[SwitchC] interface gigabitethernet 0/0/2
[SwitchC-GigabitEthernet0/0/2] stp edged-port enable
[SwitchC-GigabitEthernet0/0/2] stp bpdu-filter enable
[SwitchC-GigabitEthernet0/0/2] quit
```

在4台交换机上全局使能STP。同样因为4台交换机上的使能方法的配置方法完全一样，在此仅以SwitchA上的配置为例进行介绍。

```
[SwitchA] stp enable
```

（5）在SwitchA上配置根保护功能，即在Switch的两个指定端口上启用根保护功能，使SwitchA总为根桥。

[SwitchA] interface gigabitethernet 1/0/1

[SwitchA-GigabitEthernet1/0/1] stp root-protection

[SwitchA-GigabitEthernet1/0/1] quit

[SwitchA] interface gigabitethernet 1/0/2

[SwitchA-GigabitEthernet1/0/2] stp root-protection

[SwitchA-GigabitEthernet1/0/2] quit

以上配置完成后，过段时间，在网络计算稳定后执行以下命令，以验证配置结果。

在SwitchA上执行display stp brief命令查看端口状态和端口的保护类型，结果如下。从中可以看出将SwitchA 配置为根桥后，与 SwitchB、SwitchD 相连的端口GigabitEthernet1/0/2 和 GigabitEthernet1/0/1 在生成树计算中被选举为指定端口，并在指定端口上配置根保护功能，使得SwitchA总为根桥。

[SwitchA] display stp brief

MSTID	Port	Role	STP State	Protection
0	GigabitEthernet1/0/1	DESI	FORWARDING	ROOT
0	GigabitEthernet1/0/2	DESI	FORWARDING	ROOT

可在 SwitchB上执行 display stp interfacegigabitethernet 0/0/1 brief命令查看端口GigabitEthernet0/0/1 状态来验证，结果如下。从中可以看出 GE0/0/1 端口在生成树选举中已成为指定端口，处于Forwarding状态。

[SwitchB] display stp interfacegigabitethernet0/0/1 brief

MSTID	Port	Role	STP State	Protection
0	GigabitEthernet0/0/1	DESI	FORWARDING	NONE

可在SwitchC上执行display stp brief命令查看端口状态，结果如下。从中可以看出GE0/0/3端口在生成树选举中成为根端口，处于Forwarding状态，而GE0/0/1端口在生成树选举中成为Alternate端口，处于Discarding状态。

[SwitchC] display stp brief

MSTID	Port	Role	STP State	Protection
0	GigabitEthernet0/0/1	ALTE	DISCARDING	NONE
0	GigabitEthernet0/0/3	ROOT	FORWARDING	NONE

通过以上查看操作就可以验证以上配置是正确、成功的。

8.5 MSTP基础

通过前面的学习我们已经发现，无论是STP，还是RSTP，它们都是针对一个完整的交换网络来计算单一生成树的（所以它们都为单生成树）。这对于一些小型网络是有效的，而且配置也非常简单。但是对于一些规模比较大，结构比较复杂，特别是多 VLAN 的交换网络来说，显然会使生成树的计算更复杂，甚至无法最终形成一棵无环路的生成树。这时就得用到本节将要介绍的MSTP（Multiple Spanning Tree Protocol，多生成树协议）了。

说明

MSTP与RSTP在许多方面是完全一样的，包括主要的5种端口角色、三种端口状态、三种收敛机制、三种定时器，以及影响拓扑收敛的参数配置等，主要区别就在于MSTP可以在一个交换网络中划分多个

MST（多生成树）域，在一个MST域中又可以有多个MSTI（多生成树实例）。所以，总体上来说，MSTP的基本配置就像RSTP一样简单，不同的只是与多MST、多MSTI相关的特性了。

8.5.1 MSTP产生的背景

通过本章前面的学习已经知道，RSTP已在STP基础上进行了改进，实现了网络拓扑的快速收敛。但RSTP和STP还存在同一个缺陷：由于局域网内所有的VLAN共享一棵生成树，因此无法在VLAN间实现数据流量的负载均衡，被阻塞的冗余链路将不承载任何流量，造成带宽浪费，还有可能造成部分VLAN的报文无法转发。

在如图8-24所示网络中，如果在局域网内应用STP或RSTP，生成树结构在图中用虚线表示，S6为根交换设备。S2和S5之间、S1和S4之间的链路被阻塞，除了图中标注了“VLAN2”或“VLAN3”的链路允许对应的VLAN报文通过外，其他链路均不允 许VLAN2、VLAN3的报文通过。另外，HostA和HostB同属于VLAN2，由于S1和S4之间，以及S2和S5之间的链路被阻塞，S3和S6之间的链路又不允许VLAN2的报文通过，因此HostA和HostB之间无法互相通信。

为了弥补STP和RSTP的缺陷，IEEE于2002年发布的802.1S标准定义了MSTP。MSTP兼容STP和RSTP，既可以快速收敛，又能使不同VLAN的流量沿各自的路径转发，从而为冗余链路提供了更好的负载分担机制。

MSTP 通过把一个交换网络划分成多个域，每个域内单独形成一棵生成树，整个交换网络就可形成多棵互不影响的生成树。在MSTP中，每棵生成树叫做一个多生成树实例MSTI（Multiple Spanning Tree Instance），每个域叫做一个MST域（MST Region：Multiple Spanning Tree Region）。

MSTP 把一个生成树网络划分成多个域，每个域内形成多棵内部生成树，各个生成树实例之间彼此独立。然后，MSTP通过VLAN-生成树实例映射表把VLAN和生成树实例联系起来，将多个VLAN捆绑到一个实例中，并以实例为基础实现负载均衡。

说明

所谓实例就是一棵生成树中所包含的交换网段。通过将多个VLAN捆绑到一个实例，可以节省通信开销和资源占用率。MSTP 各个实例拓扑的计算相互独立，在这些实例上可以实现负载均衡。可以把多个相同拓扑结构的 VLAN 映射到一个实例里，这些VLAN在端口上的转发状态取决于端口在对应MSTP实例的状态。

同样以图8-24为例进行介绍，如果网络中各交换机都运行的是MSTP，就可以完全解答前面说到的STP和RSTP在本示例中的问题了。在这里可以生成以下两棵生成树，即把网络中的各 VLAN 划分到两个 MSTI 中。每个VLAN只能对应一个MSTI，即同一VLAN的数据只能在一个MSTI中传输，而一个MSTI可能对应多个VLAN。但是一个交换机可以位于多个MSTI中。

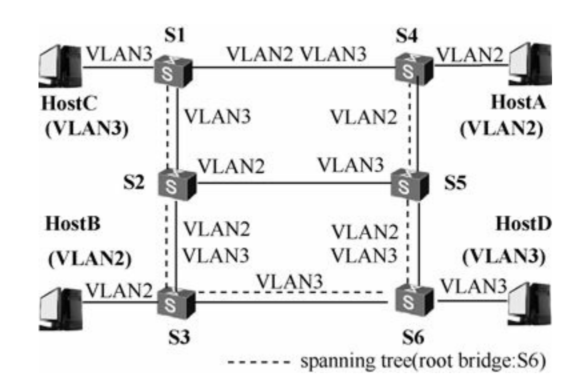


图8-24 采用STP/RSTP协议时的单生成树

（1）MSTI1：以 S4 为根桥（非根桥包括S5、S2、S3），转发VLAN2的报文。

（2）MSTI2：以 S6 为根桥（非根桥包括S3、S2、S1），转发VLAN3的报文。

S1 与 S4 之间的链路仍是阻塞的，新增了S5与S6之间的链路阻塞。这样所有 VLAN 内部可以互通，同时不同VLAN 的报文沿不同的路径转发，实现了负载分担，如图 8-25 中的 S3～S6 链路负责VLAN3报文的转发，而S2～S5链路负责VLAN2报文的转发。

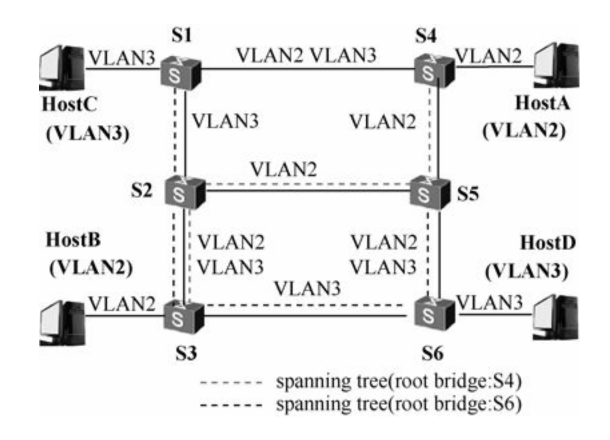


图8-25 采用MSTP协议后的两棵生成树

8.5.2 MSTP基本概念

因为在MSTP网络中可以有多棵生成树实例，就涉及到生成树实例的划分及各生成树实例之间的关系等问题，所以与单生成树的STP和RSTP在许多方面存在不同。本节具体介绍MSTP所涉及的一些基本概念。

1．MSTP网络的层次结构

MSTP不仅涉及多个MSTI（生成树实例），而且还可划分多个MST域（MST Region，也称为MST区域）。总的来说，一个MSTP网络可以包含一个或多个MST域，而每个MST 域中又可包含一个或多个MSTI。组成每个 MSTI 的是其中运行 STP/RSTP/MSTP的交换设备，是这些交换设备经MSTP协议计算后形成的树状网络。

如图8-26所示的MSTP网络中划分了3个MST区域，每个区域中又包括了3个MSTI。

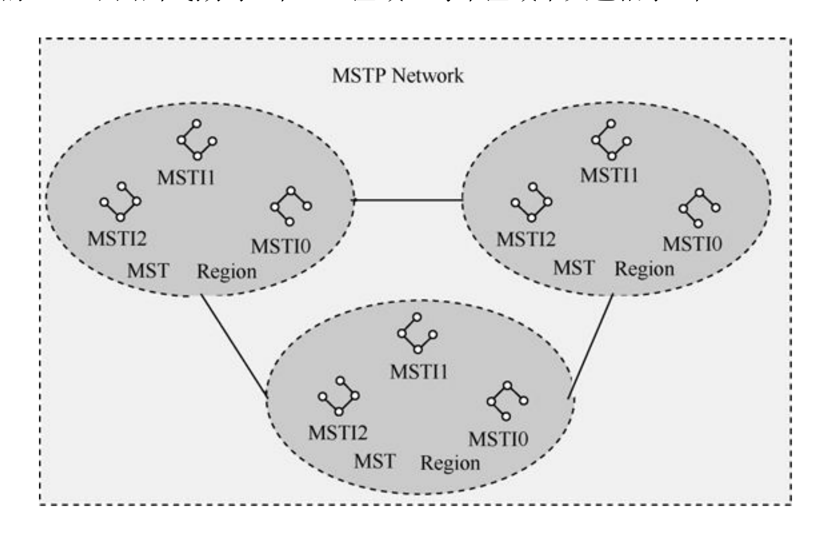


图8-26 MSTP网络示例

2．MST域

MST域（Multiple Spanning Tree Region，多生成树域），由交换网络中的多台交换设备以及它们之间的网段所构成。同一个MST域的设备具有下列特点。

（1）都启动了MSTP。

（2）具有相同的域名。

（3）具有相同的VLAN到生成树实例映射配置。

（4）具有相同的MSTP修订级别配置。

一个MSTP网络可以存在多个MST域，各MST域之间在物理上直接或间接相连。用户可以通过MSTP配置命令把多台交换设备划分在同一个MST域内。

图8-27所示的MST域D0中是由交换机S1、S2、S3和S4构成，域中有3个MSTI，即MSTI0、MSTI1和MSTI2。

3．MSTI

MSTI（Multiple Spanning Tree Instance，多生成树实例）是指MST域内的生成树。一个MST域内可以通过MSTP生成多棵生成树，各棵生成树之间彼此独立。一个MSTI可以与一个或者多个VLAN对应，但一个VLAN只能与一个MSTI对应。

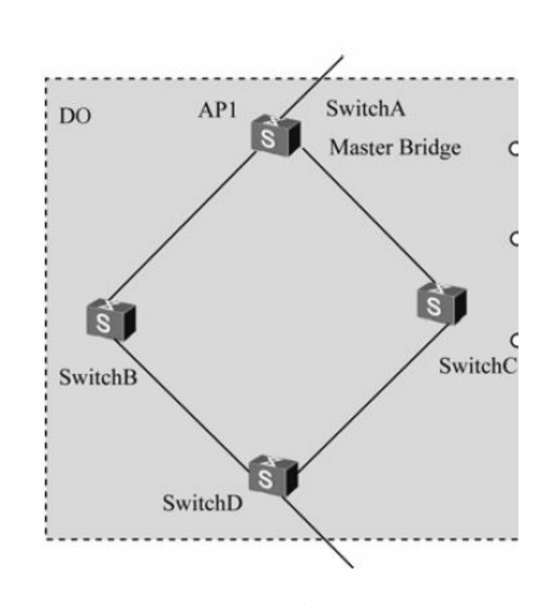


图8-27 MST域示例

既然是生成树，那就不允许存在环路。仍以图8-27所示的MSTP网络为例进行介绍，这个MST域中包括了3个MSTI（图中的MSTI0、MSTI1、MSTI2，注意看它们的拓扑，总有一个方向的交换机连接是断开的），每个MSTI都没有环路。

再看一下图8-28所示的示例。在这个MST域的交换网络中包括了3个VLAN：VLAN 10、VLAN 20和VLAN 30。这时又该划分成多少个MSTI呢？如果我们把VLAN 10和VLAN 20放进一个MSTI中，则所得到的拓扑如图 8-29的左图所示，明显存在环路；如果把VLAN 10和VLAN 30放进一个MSTI中，得到如图 8-29所示的拓扑，也明显存在环路；同样如果把VLAN 20和VLAN 30划分到一个MSTI中，则拓扑如图 8-29的右图所示，也存在环路。这时只好为每个 VLAN 单独划分成一个 MSTI，得到的每个 MSTI拓扑如图8-30所示

的3个MSTI，就不存在环路了（注意，虚线所代表的是通过MSTP协议配置阻塞的链路），确保每个MSTI中没有环路出现。

在一般的企业网络中，通常是将支持MSTP的设备全部划分到一个MST域中，而将不支持MSTP的设备划分到另一个MST域中。对于MSTI来说，通常是将具有相同转发路径的VLAN映射到一个MSTI中，以形成一棵独立的生成树。

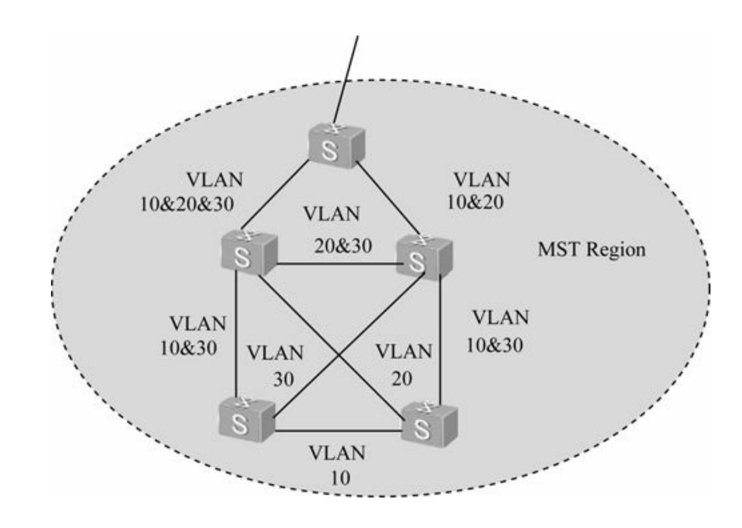


图8-28 MSTI划分示例

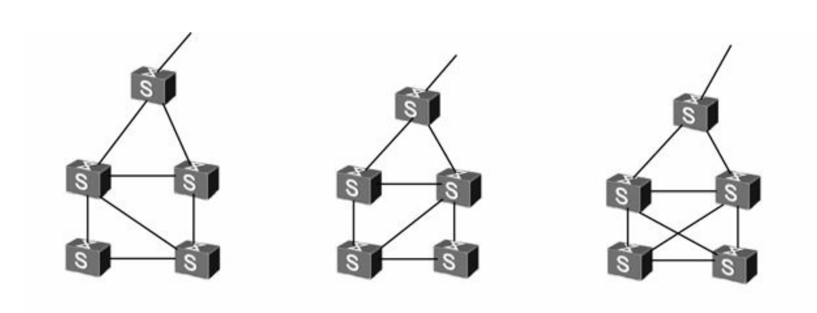


图8-29 每个MSTI放进两个VLAN情况下的生成树拓扑

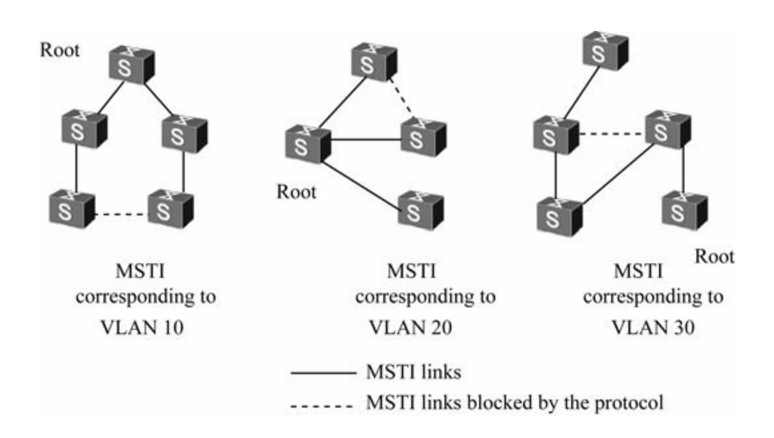


图8-30 每个MSTI对应一个VLAN的生成树拓扑

4．VLAN映射表

VLAN映射表是MST域的属性，描述了VLAN和MST域中对应MSTI之间的映射关系。也就是把那些VLAN 分别加入哪个 MSTI中。一个 VLAN 只能加入一个 MSTI中，即同一VLAN的数据只能在一个MSTI中传输，而一个MSTI可能对应多个VLAN。但是一台交换机可以位于多个MSTI中，毕竟一台交换机上可以划分多个VLAN。

如图8-27所示的MSTP网络，MST域D0中所包括的VLAN映射表如下。

（1）VLAN1映射到MSTI1。

（2）VLAN2和VLAN3映射到MSTI2。

（3）其余VLAN映射到MSTI0。

5．IST

IST（Internal Spanning Tree，内部生成树）是各个MST域内部的一棵生成树，是仅针对具体的**MST**域来计算的 。但它是一个特殊的MSTI，其MSTI ID为 0，即 IST通常称为MSTI0。每个**MST**域中只有一个**IST**，包括对应MST域中所有互联的交换机。

在如图8-31所示的MSTP网络中（包括了多个MST域）每个MST域内部用细线连接的各交换机就构成了对应MST域中的IST。

6．CST

CST（Common Spanning Tree，公共生成树）是连接整个MSTP网络内所有MST域的一棵单生成树，是针对整个**MSTP**网络来计算的 。如果把每个MST域看作是一台“交换机”，每个MST域看成CST的一个节点，则CST就是这些节点“交换机”通过STP 或者RSTP协议计算生成的一棵生成树（SST）。即每个**MSTP**网络中只有一个**CST** 。每个MST域中的IST是整个MSTP网络CIST在对应MST域中的一个片段。

在图8-31中用于连接各个MST域的粗线条连接就构成了CST。

7．CIST

CIST（Common and Internal Spanning Tree，公共和内部生成树）是通过STP 或RSTP协议计算生成的，连接整个MSTP 网络内所有交换机的单生成树，由**IST** 和**CST** 共同构成 。这里要注意了，上面介绍的CST是连接交换网络中所有**MST** 域 的单生成树，而此处的CIST则是连接交换网络内的所有交换机 的单生成树。即每个**MSTP** 网络中也只有一个**CIST** 。交换网络中的所有MST 域的IST 和CST 一起构成一棵完整的生成树，也就是这里的CIST。

在图8-31中，A0、B0、C0、D0四个MST区域中的IST，再加上MST域间的CST就是整个交换网络的CIST了。

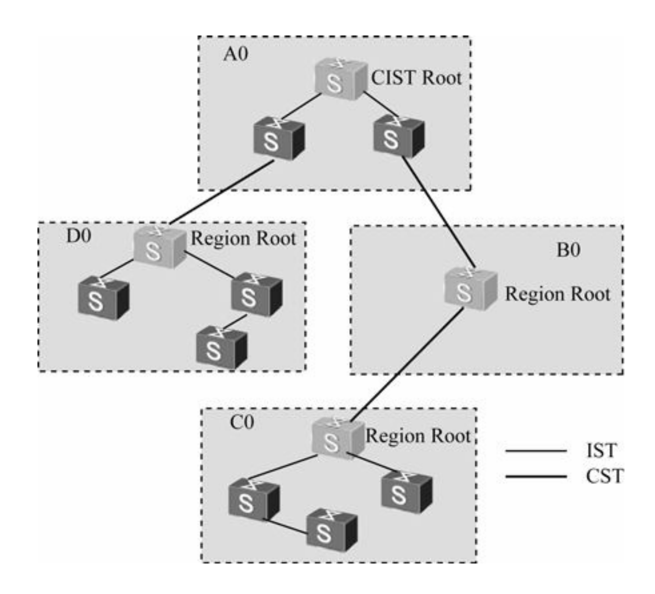


图8-31 多MST域的MSTP网络示例

8．SST

构成SST（Single Spanning Tree，单生成树）有两种情况。

（1）运行STP或RSTP生成树协议的交换机只属于一个生成树。

（2）MST域中只有一个交换机，这个交换机构成单生成树。

在图 8-31 中，B0 域中的交换机就是一棵单生成树，因为此域中只有一台交换机。

9．总根

总根是CIST生成树的根桥，通常是交换网络中最上层的交换机。图8-31中总根是在A0域中IST生成树的根桥。一个MSTP网络只有一个总根。

10．域根

因为在MSTP网络中，每MST域都有一个特殊的IST实例，以及许多MSTI实例，所以域根（Regional Root）又分为 IST域根和MSTI域根。

各个MST域中的IST生成树中距离CIST总根最近的交换机是IST域根。总根所在MST 域的 IST 域根就是总根。在图 8-31 的示例中，也已标出了非总根所在的 B0、C0 和D0三个MST域的IST域根。

MSTI 的域根是对应生成树实例的树根，域中不同的 MSTI 有各自的域根。而且， MST域内各棵生成树的拓扑不同，域根也可能不同。

8.5.3 MSTP的端口角色

MSTP中的端口角色主要有根端口（root port）、指定端口（designated port）、替代端口（alternate port）、备份端口（backup port）、主端口（master port）、域边缘端口和边缘端口。其中，根端口、指定端口、Alternate端口、Backup端口和边缘端口这五种主要端口角色的作用与RSTP协议中对应的端口角色定义完全相同。

与RSTP一样，在MSTP中也是除边缘端口外，其他端口角色都参与MSTP的计算过程。而且，同一端口在不同的生成树实例中可以担任不同的角色。为了便于说明，下面给出一个典型的MSTP端口示例，如图8-32所示。

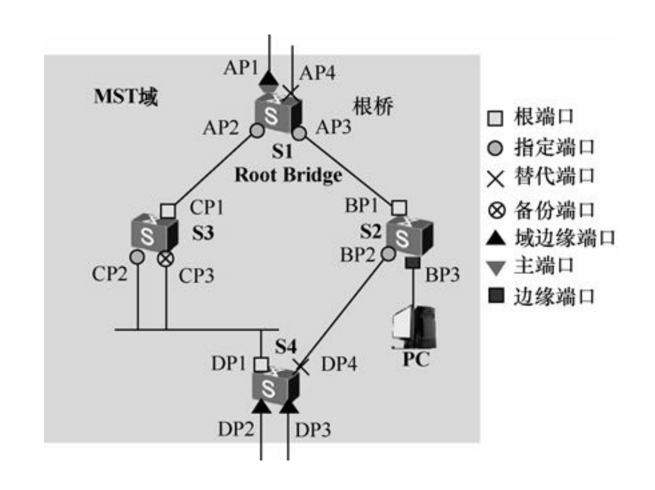


图8-32 MSTP端口示例

1．根端口

根端口仅针对非根桥而言，非根桥上到根桥距离开销最小的端口就是本交换机的根端口。在到根桥距离开销相同的情况下，离根桥最近的端口是本交换机的根端口。根端口负责向树根方向转发数据。根桥上没有根端口，只有下面将要介绍指定端口。

在图8-32中，S1为根桥，CP1为S3的根端口，BP1为S2的根端口，DP1为S4的根端口。

2．指定端口

对一台交换机而言，它的指定端口是向下游交换机转发BPDU报文的端口。交换机连接下级交换机的所有端口都是指定端口，不仅根桥上有，非根桥上同样有。

在图8-32中，AP2和AP3和BP2分别为S1和S2的指定端口，CP2为S3的指定端口。

3．边缘端口

如果指定端口位于整个域的边缘，不再与任何交换机连接，这种端口叫做边缘端口。边缘端口一般与用户终端设备（如PC机）直接连接。

在图8-32中，BP3为边缘端口。

4．Alternate端口

从发送BPDU来看，Alternate端口就是由于学习到其他交换机发送的BPDU而被阻塞的端口。从转发用户流量来看，Alternate端口提供了从指定桥到根桥的一条备份路径，所以Alternate端口是根端口的备份端口，如果根端口被阻塞后，Alternate端口将成为新的根端口。

在图8-32中，DP4和AP4为Alternate端口。

5．Backup端口

当同一台交换机的两个端口同时连接一个设备时就存在一个环路，此时交换机会将其中一个端口阻塞，这个端口就是Backup端口。

从发送 BPDU 来看，Backup 端口就是由于学习到本设备上其他端口发送的 BPDU而被阻塞的端口。从转发用户流量来看，Backup端口，作为指定端口的备份，提供了一条从根桥到下级设备的备份通路。

在图8-32中，CP3为Backup端口，因为它与CP2端口同时连接到下游的同一个设备。

6．Master端口

Master 端口是 MST 域和总根相连的所有路径中最短路径上的端口，它是交换机上连接 MST 域到总根的端口。Master端口是域中的报文去往总根的必经之路。Master 端口是特殊域边缘端口，Master端口在CST/CIST上的角色是根端口，在其他各实例上的角色都是Master。

在图8-32中，交换设备S1、S2、S3、S4和它们之间的链路构成一个MST域，S1交换设备的端口AP1在域内的所有端口中到总根的路径开销最小，所以AP1为Master端口。

7．域边缘端口

域边缘端口是指位于MST域的边缘并连接其他MST域或SST的端口。进行MSTP计算时，域边缘端口在MSTI 上的角色和 CIST 实例的角色保持一致。即如果边缘端口在 CIST 实例上的角色是 Master 端口（连接域到总根的端口），则它在域内所有 MSTI上的角色也是Master端口。

在图8-32中，AP1是域边缘端口，它在CIST上的角色是Master端口，则AP1在MST域内所有生成树实例上的角色都是Master端口。

8.5.4 MSTP的端口状态与收敛机制

MSTP定义的端口状态也与RSTP协议中的定义完全相同，也是根据端口是否转发用户流量、接收/发送BPDU报文，把端口状态划分为3种。

（1）Forwarding状态：转发状态，既转发用户流量又接收/发送BPDU报文。

（2）Learning状态：学习状态，不转发用户流量，只接收/发送BPDU报文。

（3）Discarding状态：丢弃状态，只接收BPDU报文，不转发报文。

与RSTP中的端口状态一样，MSTP的端口状态和端口角色是没有必然联系的，表8-11给出了各种端口角色能够具有的端口状态。

表8-11 MSTP各种端口角色具有的端口状态

端口角色 端口状态	根端口/Master 端口	指定端口	域边缘端口	Alternate 端口	Backup 端口
Forwarding	√	√	√	—	—
Learning	√	√	√	—	—
Discarding	√	√	√	√	√

说明

MSTP的收敛机制与RSTP是完全一样的，具体参见本章8.3.4和8.3.5节。在P/A 机制方面同样支持普通模式和增强模式两种，具体参见本章8.4.7节。

8.5.5 MSTP拓扑计算原理

MSTP将整个二层网络划分为多个MST域，把每个域视为一个节点。各个**MST**域之间按照**STP**或者**RSTP**协议算法进行计算并生成**CST**（是单生成树）；在一个**MST**域内则是通过**MSTP**协议算法计算生成若干个**MSTI**（是多生成树），其中实例0被称为 IST。MSTP使用MST BPDU（Multiple Spanning Tree Bridge Protocol Data Unit，多生成树桥协议数据单元）作为生成树计算的依据。MST BPDU报文用来计算生成树的拓扑、维护网络拓扑以及传达拓扑变化记录。

1．MSTP向量优先级

MSTI 和 CIST 拓扑都是根据优先级向量来计算的，这些优先级向量信息都包含在MST BPDU中。各交换机互相交换MST BPDU来生成MSTI和CIST。

参与CIST计算的优先级向量按优先级别从高到低依次是根桥ID、外部路径开销、域根ID、内部路径开销、指定桥ID、指定端口ID、接收端口ID；参与MSTI计算的优先级向量按优先级别从高到低依次是域根ID、内部路径开销、指定桥ID、指定端口ID、接收端口ID。

以上这些优先级向量说明如表8-12所示。

表8-12 优先级向量说明

优先级向量名	说明
根桥 ID	根桥 ID 用于选择 CIST 中的根桥。在 BPDU 中对应的网桥 ID，计算公式为：Priority(16bits)+MAC(48bits)
外部路径开销（ERPC）	从 MST 域根到达总根的路径开销。MST 域内所有交换机上保存的外部路径开销相同。若 CIST 根桥在域中，则域内所有交换机上保存的外部路径开销为 0
域根 ID	也就是通常所说的 MSTI 树根，域根 ID 用于选择 MSTI 中的树根。它也是通过网桥 ID 来选举的，计算公式为：Priority(16bits)+MAC(48bits)
内部路径开销（IRPC）	本交换机到达域根桥的路径开销。域边缘端口保存的内部路径开销值大于（优先级越低）非域边缘端口保存的内部路径开销
指定桥	CIST 或 MSTI 实例的指定桥是本交换机通往域根的最邻近的上游交换机。如果本交换机就是总根或域根，则指定桥为自己
指定端口	指定桥上与本交换机根端口相连的端口就是指定端口。其端口 ID（Port ID）= Priority(8 位)＋ 端口号(8 位)。端口优先级必须是 16 的整数倍
接收端口	接收到 BPDU 报文的端口。其端口 ID （Port ID）= Priority(8 位)＋ 端口号(8 位)。端口优先级必须是 16 的整数倍

同一类向量比较时，值最小的向量具有最高优先级。具体比较规则如下。

（1）首先，比较根桥ID。

（2）如果根桥ID相同，再比较外部路径开销。

（3）如果外部路径开销还相同，再比较域根ID。

（4）如果域根ID仍然相同，再比较内部路径开销。

（5）如果内部路径仍然相同，再比较指定桥ID。

（6）如果指定桥ID仍然相同，再比较指定端口ID。

（7）如果指定端口ID还相同，再比较接收端口ID。

如果端口接收到的BPDU内包含的配置消息优于端口上保存的配置消息，则端口上原来保存的配置消息被新收到的配置消息替代。端口同时更新交换机保存的全局配置消息。反之，新收到的BPDU被丢弃。

2．CIST的计算

经过配置消息比较后，首先在整个网络中选择一个优先级最高的交换机作为 CIST的树根，然后在每个MST域内通过MSTP协议算法计算生成IST；同时MSTP将每个MST域作为单台交换机对待，通过STP或者RSTP协议算法在MST域间计算生成CST。CST和IST构成了整个交换机网络的CIST。

3．MSTI的计算

在MST域内，MSTP根据VLAN和生成树实例的映射关系，针对不同的VLAN生成不同的生成树实例。MSTI具有以下的特点。

（1）每个MSTI独立计算自己的生成树，互不干扰。

（2）每个MSTI的生成树计算方法与RSTP基本相同。

（3）每个MSTI的生成树可以有不同的根，不同的拓扑。

（4）每个MSTI在自己的生成树内发送BPDU。

（5）每个MSTI的拓扑通过命令配置决定（不是自动生成的）。

（6）每个端口在不同MSTI上的生成树参数可以不同。

（7）每个端口在不同MSTI上的角色、状态可以不同。

4．MSTI生成树算法实现

在一开始时，每台交换机的各个端口会生成以自身交换机为根桥的配置消息，其中根路径开销为 0，指定桥 ID 为自身交换机 ID，指定端口为本端口。每台交换机都向外发送自己的配置消息，并在接收到其他配置消息后进行如下处理。

（1）当端口收到比自身的配置消息优先级低（优先级的比较就是根据前面介绍的向量优先级比较规则

进行的）的配置消息时，交换机把接收到的配置消息丢弃，对该端口的配置消息不作任何处理。

（2）当端口收到比本端口配置消息优先级高的配置消息时，交换机把接收到的配置消息中的内容替换该端口的配置消息中的内容；然后交换机将该端口的配置消息和交换机上的其他端口的配置消息进行比较，选出最优的配置消息。

计算生成树的步骤如下。

（1）选举根桥。此步是通过比较所有交换机发送的配置消息的树根ID，树根ID值最小的交换机为CIST根桥，或者MST域根桥。

（2）选举非根桥上的根端口。每台非根桥把接收到最优配置消息的那个端口定为自 身交换机的根端口。

（3）选举指定端口。在这一步又分为以下两个子步骤：

首先，交换机根据根端口的配置消息和根端口的路径开销，为每个端口计算一个标准的指定端口配置消息：用树根ID替换为根端口配置消息中的树根ID；用根路径开销替换为根端口配置消息中的根路径开销加上根端口的路径开销；用指定桥 ID 替换为自身交换机的ID；用指定端口ID替换为自身端口ID。

然后，交换机对以上规则计算出来的配置消息和对应端口上原来的配置消息进行比较。如果端口上原来的配置消息更优，则交换机将此端口阻塞，端口的配置消息不变，并且此端口将不再转发数据，只接收配置消息（相当于根端口）；如果通过以上替换计算出来的配置消息比端口上原来的配置消息更优，则交换机就将该端口设置为指定端口，端口上的配置消息替换成通过以上替换计算出来的配置消息，并周期性向外发送。

（4）在MSTI生成树拓扑收敛后，无论非根桥是否接收到根桥传来的信息都按照Hello定时器周期性发送BPDU。如果一个端口连续3个Hello时间（这个是缺省的设置）接收不到指定桥（也就是它所连接的上一级交换机）送来的BPDU，那么该交换机认为与此邻居之间的链路失败。

5．MSTP对拓扑变化的处理

在MSTP中检测拓扑是否发生了变化的标准是根据一个非边缘端口的状态是否迁移到Forwarding状态，如果是迁移到了Forwarding状态，则会发生拓扑变化。

交换机一旦检测到拓扑发生变化，进行如下处理。

（1）为本交换机的所有非边缘指定端口启动一个 TC While Timer（该计时器值是Hello Time的两倍），并在这个时间内，清空这些端口上学来的MAC地址。如果是根端口上有状态变化，则启动根端口。

（2）发生状态变化的这些端口向外发送TC BPDU，其中的TC置位，直到TC While Timer超时。根端口总是要发送这种TC BPDU。

其他交换机接收到TC BPDU，进行如下处理。

（1）清空所有端口学来的MAC地址，收到TC BPDU的端口除外。

（2）为所有自己的非边缘指定端口和自己的根端口启动TC While计时器，重复上述过程。

8.5.6 MSTP BPDU报文

MSTP使用MST BPDU（Multiple Spanning Tree Bridge Protocol Data Unit，多生成树桥协议数据单元）作为生成树计算的依据。MST BPDU报文用来计算生成树的拓扑、维护网络拓扑以及传达拓扑变化记录。

STP中定义的配置BPDU、RSTP中定义的RST BPDU、MSTP中定义的MST BPDU及TCN BPDU在版本号和类型值方面的比较如表 8-13所示。

表8-13 4种BPDU的比较

名称	版本	类型
配置 BPDU	0	0x00
RST BPDU	2	0x02
MST BPDU	3	0x02
TCN BPDU	0	0x80

1．MSTP BPDU报文格式

MST BPDU报文结构如图 8-33所示。无论是域内的MST BPDU还是域间的MST BPDU，前 36个字节和RST BPDU相同。从第 37个字节开始是MSTP专有字段。最后的MSTI配置信息字段由若干MSTI配置信息组连缀而成。各字段说明如表8-14所示。其实这里的CIST相当于RSTP中的单生成树。

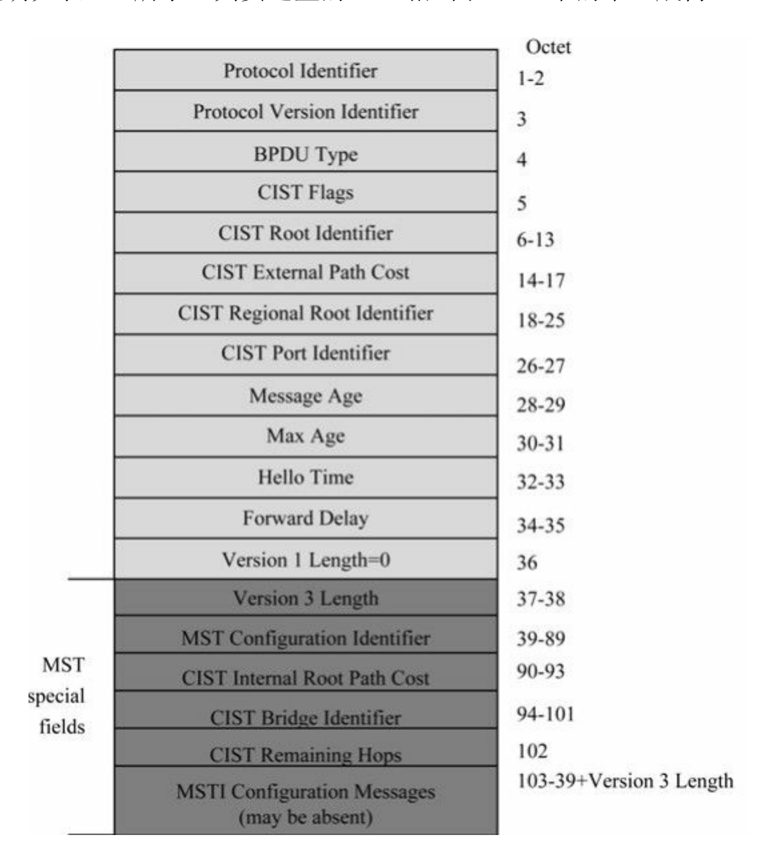


图8-33 MST BPDU格式

表8-14 MST BPDU格式字段说明

字段内容	字节数	说明
Protocol Identifier	2	协议标识符，目前总为 0
Protocol Version Identifier	1	协议版本标识符，STP 为 0，RSTP 为 2，MSTP 为 3
BPDU Type	1	BPDU 类型： • 0x00：STP 的 Configuration BPDU • 0x80：STP 的 TCN BPDU • 0x02：RST BPDU 或者 MST BPDU
CIST Flags	1	CIST 标志字段，与 RSTP 中的标志字段完全一样
CIST Root Identifier	8	CIST 的总根桥 ID
CIST External Path Cost	4	CIST 外部路径开销，指从本桥所属的 MST 域到 CIST 根桥所属的 MST 域的累计路径开销，类似于 RSTP 中的根路径开销，也是根据链路带宽计算的
CIST Regional Root Identifier	8	CIST 的域根桥 ID，即 IST Master 的 ID。如果总根在这个域内，那么该域的根桥 ID 就是总根桥 ID

（续表）

字段内容	字节数	说明
CIST Port Identifier	2	发送 BPDU 报文的端口在 IST 中的指定端口 ID
Message Age	2	MST BPDU 报文的生存期
Max Age	2	MST BPDU 报文的最大生存期，超时则认为到根交换设备的链路故障
Hello Time	2	Hello 定时器，缺省为 2s
Forward Delay	2	Forward Delay 定时器，缺省为 15s
Version 1 Length	1	Version1 BPDU 的长度，值固定为 0
Version 3 Length	2	Version3 BPDU 的长度
MST Configuration Identifier	51	MST 配置标识符，表示 MST 域的标签信息，包含 4 个字段，如图 8-34 所示。只有这里面的 4 个字段完全相同的，并且互联的交换设备，才属于同一个域。这四字段的说明如表 8-15 所示
CIST Internal Root Path Cost	4	CIST 内部路径开销，指从发送 BPDU 报文的端口到 IST Master（主桥）的累计路径开销。CIST 内部路径开销也是根据链路带宽计算的
CIST Bridge Identifier	8	CIST 的指定桥 ID
CIST Remaining Hops	1	BPDU 报文在 CIST 中的剩余跳数(每经过一个桥设备跳数减 1）
MSTI Configuration Messages(may be absent)	16	MSTI 配置消息。每个 MSTI 的配置消息占 16 个字节，如果有 *n* 个 MSTI 就占用 *n*×16bytes。单个 MSTI 配置消息的结构如图 8-35 所示，字段说明如表 8-16 所示

字段	Octet
Configuration Identifier Format Selector	39
Configuration Name	40-71
Revision Level	72-73
Configuration Digest	74-89

图8-34 MST配置标识符结构

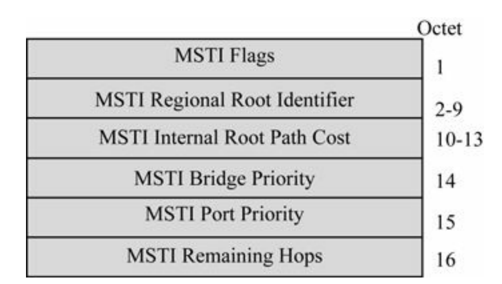


图8-35 MSTI配置消息结构

表8-15 MST配置标识符字段说明

字段	字节数	说明
Configuration Identifier Format Selector	1	配置标识符格式选择器，固定为 0
Configuration Name	32	MST 域名，32 字节长字符串，每个 MST 域有唯一的配置消息
Revision Level	2	MST 配置修订级别，2 字节非负整数
Configuration Digest	16	配置摘要，利用 HMAC-MD5 算法将域中 VLAN 和实例的映射关系加密成 16 字节的摘要

表8-16 MSTI配置消息字段说明

字段	字节数	说明
MSTI Flags	1	MSTI 标志位
MSTI Regional Root Identifier	8	MSTI 域根桥 ID

（续表）

字段	字节数	说明
MSTI Internal Root Path Cost	4	MSTI 内部路径开销，指从本端口到 MSTI 域根桥的累计路径开销。MSTI 内部路径开销根据链路带宽计算
MSTI Bridge Priority	1	本桥在 MSTI 中的桥优先级
MSTI Port Priority	1	发送 MST BPDU 的端口在 MSTI 中的端口优先级
MSTI Remaining Hops	1	BPDU 报文在 MSTI 中的剩余跳数

2．MSTP BPDU报文格式可配置功能

目前MSTP的BPDU报文存在两种格式。

（1）dot1s：IEEE802.1s规定的报文格式。

（2）legacy：华为私有协议报文格式。

如果端口收发报文格式为缺省支持 dot1s 或者 legacy，这样就存在一个缺点：需要人工识别对端的BPDU报文格式，然后手工配置命令来决定支持哪种格式。人工识别报文格式比较困难，且一旦配置错误，就有可能导致MSTP计算错误，出现环路。

华为技术有限公司采用的端口收发MSTP报文格式可配置（stp compliance）功能，支持自动识别（auto）模式，这样就能够实现对 BPDU报文格式的自适应。这样报文收发不但支持dot1s和legacy格式，还能通过auto模式根据收到的BPDU报文格式自动切换接口支持的 BPDU 报文格式，使报文格式与对端匹配。在自适应的情况下，接口初始支持dot1s格式，收到报文后，格式则和收到的报文格式保持一致。

3．每个Hello Time时间内端口最多能发送BPDU的报文数可配置功能

Hello Time定时器用于生成树协议定时发送配置消息维护生成树的稳定。如果交换设备在一段时间内没有收到BPDU报文，则会由于消息超时而对生成树进行重新计算。当交换设备成为根交换设备时，该交换设备会以该设置值为时间间隔发送BPDU报文。非根交换设备采用根交换设备所设置的Hello Time时间值。

华为 S系列交换机提供的每个Hello Time时间内端口最多能够发送的BPDU报文个数可配置（Max Transmitted BPDU Number in Hello Time is Configurable）功能，可以设定当前端口在Hello Time时间内配置BPDU的最大发送数目。用户配置的数值越大，表示每Hello Time时间内发送的报文数越多。适当地设置该值可以限制端口每Hello Time时间内能发送的BPDU数目，防止在网络拓扑动荡时，BPDU占用过多的带宽资源。

8.6 MSTP配置

在了解了MSTP的一些主要基础知识和工作原理后，本节就要介绍MSTP协议的具体配置与管理方法了。下面先同样了解一下华为S系列交换机的MSTP相关参数缺省配置，如表8-17所示，实际应用的配置可以基于缺省配置进行修改。

表8-17 MSTP相关参数的缺省配置

参数	缺省值
生成树协议工作模式	MSTP 模式
MSTP 功能	全局 MSTP 功能使能，接口的 MSTP 功能使能
交换设备的优先级	32768
端口的优先级	128
路径开销缺省值的计算方法	dot1t，即 IEEE 802.1t 标准
Forward Delay Time	1 500 厘秒
Hello Time	200 厘秒
Max Age Time	2 000 厘秒

8.6.1 MSTP基本功能主要配置任务

MSTP 可以把一个交换网络划分成多个域，每个域内形成多棵生成树，生成树之间彼此独立，实现不同VLAN流量的分离，达到网络负载均衡的目的。

通过给交换设备配置MSTP的工作模式、配置域并激活后，启动MSTP，MSTP便开始进行生成树计算，将网络修剪成树状，破除环路。但是，如果需要人为干预生成树计算的结果，还可以进行如下配置：手动配置指定根桥和备份根桥设备，配置交换设备在指定生成树实例中的优先级数值，配置端口在指定生成树实例中的路径开销数值，配置端口在指定生成树实例中的优先级数值。

下面具体介绍这些MSTP基本功能配置任务。

1．配置MSTP工作模式

这一项配置任务很简单，就是指定交换设备工作在MSTP协议下。MSTP兼容STP和RSTP。缺省情况下，交换设备的工作模式为MSTP。但要注意的是，因为STP和MSTP不能互相识别报文，而MSTP和RSTP可以互相识别报文，所以如果设备工作在MSTP工作模式下就会设置所有与运行STP的交换设备直接相连的端口工作在STP模式下，其他端口工作在MSTP模式下，实现运行不同生成树协议的设备之间的互通。

2．配置并激活MST域

MST域是由交换网络中的多台交换设备以及它们之间的网段所构成。这些交换设备启动MSTP后，具有相同域名、相同VLAN到生成树映射配置和相同MSTP修订级别配置，并且物理上直接相连。一个交换网络可以存在多个 MST 域，用户可以通过 MSTP配置命令把多台交换设备划分在同一个MST域内。

3．（可选）配置根桥和备份根桥

可以通过生成树计算来自动确定生成树的根桥，用户也可以手动配置设备为指定生成树的根桥或备份根桥。在一棵生成树中，生效的根桥只有一个；当两台或两台以上的设备被指定为同一棵生成树的根桥时，系统将选择MAC地址最小的设备作为根桥。

可以在每棵生成树中指定多个备份根桥。当根桥出现故障或被关机时，备份根桥可以取代根桥成为指定生成树的根桥；但此时如果配置了新的根桥，则备份根桥不会成为根桥；如果配置了多个备份根桥，则MAC地址最小的备份根桥将成为指定生成树的根桥。

设备在各生成树中的角色互相独立，一台交换设备在作为一棵生成树的根桥或备份根 桥的同时，也可以作为其他生成树的根桥或备份根桥；但在同一棵生成树中，一台设备不能既作为根桥，又作为备份根桥。在配置MSTP过程中，建议手动配置根桥和备份根桥。

4．（可选）配置交换设备在指定生成树实例中的优先级

在一个生成树实例中有且仅有一个根桥，它是该生成树实例的逻辑中心。在进行根桥的选择时，一般会希望选择性能高、网络层次高的交换设备作为根桥。但是，性能高、网络层次高的交换设备其优先级不一定高，因此需要配置优先级以保证该设备成为根桥。交换设备在指定生成树实例中的优先级值越小，则交换设备的优先级越高，成为该生成树实例根桥的可能性越大。

对于生成树实例中部分性能低、网络层次低的交换设备，不适合作为根桥设备，一般会配置其优先级以保证该设备不会成为根桥。

5. （可选）配置端口在指定生成树实例中的路径开销

路径开销是一个端口量，是MSTP协议用于选择链路的参考值。端口的路径开销是生成树计算的重要依据，在不同生成树实例中为同一端口配置不同的路径开销值，可以使不同VLAN的流量沿不同的物理链路转发，实现VLAN的负载分担功能。

在同一种计算方法下，端口开销值越小，端口在该生成树实例中到根桥的路径开销越小，成为根端口的可能性就越大。端口路径开销会影响指定生成树实例中根端口的选择，在该实例中某台设备所有端口到达根桥路径开销最小者，就是根端口。在存在环路的网络环境中，对于链路速率值相对较小的端口，建议将其路径开销值配置相对较大，以使其在生成树算法中被选举成为阻塞端口，阻塞其所在链路。

6. （可选）配置端口在指定生成树实例中的优先级

在参与MSTP生成树计算时，对于处在生成树实例中的交换设备端口，其优先级的高低会影响到是否被选举为指定端口。端口优先级值越小，端口在该生成树实例中成为指定端口的可能性就越大；值越大，端口在该生成树实例中成为指定端口的可能性越小。如果希望将生成树实例中的某交换设备的端口阻塞从而破除环路，则可将其端口优先级值设置比缺省值大，使得在选举过程中成为被阻塞的端口。

7. 启用MSTP

当交换设备配置MSTP基本功能后，必须使能设备MSTP功能，MSTP相关配置才能生效。

在环形网络中一旦启用 MSTP，MSTP 便立即进行生成树计算。而且，诸如交换设备的优先级、端口优先级等参数都会影响到生成树的计算，在计算过程中这些参数的变动可能会导致网络振荡。为了保证生成树计算过程快速而且稳定，必须在对交换设备及其端口进行必要的基本配置以后再启用MSTP。

8. 配置收敛方式

当生成树的拓扑结构发生改变时，和它建立映射关系的VLAN的转发路径也将发生变化。此时，交换设备的ARP表中与这些VLAN相关的表项也需要更新。根据对ARP表项的处理方式不同，MSTP的收敛方式分为fast和normal两种。

（1）fast：ARP表将需要更新的表项直接删除。

（2）normal：ARP表中需要更新的表项快速老化。

交换设备将ARP表中这些表项的剩余存活时间置为0，对这些表项进行老化处理。如果配置的 ARP 老化探测次数大于零，则 ARP 对这些表项进行老化探测。建议选择 normal收敛方式。若选择fast方式，频繁的ARP表项删除可能会导致设备CPU占用率高达100%，报文处理超时导致网络振荡。

8.6.2 配置MSTP基本功能

MSTP 可阻塞二层网络中的冗余链路，将网络修剪成树状，达到消除环路的目的。同时，MSTP引入多实例，通过将不同VLAN映射到不同实例中，实现不同VLAN的流量负载分担。MSTP的基本配置思路如下。

（1）在环形网络中，划分MST域，在域中配置不同的MSTI。

（2）为各个MSTI选出其中的一个交换设备作为根桥。

（3）在各个MSTI中计算出其他交换设备到根桥的最短路径，为每个非根桥设备选举一个根端口。

（4）在各个MSTI中通过端口ID为每个连接选举出一个指定端口。

上节介绍的八项MSTP基本功能配置任务的具体配置步骤如表8-18所示。

表8-18 MSTP基本功能配置步骤

配置任务	步骤	命令	说明
配置MSTP工作模式	1	**system-view** 例如：<HUAWEI> **system-view**	进入系统视图
	2	**stp mode mrstp** 例如：[HUAWEI] **stp mode mstp**	配置交换机的 MSTP 生成树工作模式。执行本命令后，在交换设备所有启用生成树协议的端口中，除了和 STP 交换设备直接相连的端口工作在 STP 模式下，其他端口都工作在 MSTP 模式下，即向外发送 MST BPDU 报文 缺省情况下，除 S2700SI 子系列运行模式为 STP 模式外，其他系列为运行 MSTP 模式，MSTP 模式兼容 STP 和 RSTP 模式，可用 **undo stp mode** 命令恢复交换设备的缺省生成树协议工作模式
配置并激活MST域	3	**stp region-configuration** 例如：[HUAWEI] **stp region-configuration**	进入 MST 域视图，进行 MST 域配置。只要两台交换设备的以下配置相同，这两台交换设备才属于同一个 MST 域。 （1）MST 域的域名 （2）多生成树实例和 VLAN 的映射关系 （3）MST 域的修订级别 当需要为当前设备或 MSTP 进程配置上述 3 个参数时，就需要通过本命令进入 MST 域视图。缺省情况下，这 3 个参数均取以下缺省值。 （1）MST 域名为交换设备主控板的 MAC 地址 （2）MSTP 修订级别取值为 0 （3）所有 VLAN 均映射到 CIST 上 可用 **undo stp region-configuration** 命令将 MST 域配置恢复为缺省值
	4	**region-name** *name* 例如：[HUAWEI-mst-region] **region-name** lycb	配置 MST 域的域名，为 1～32 个字符，不支持空格，区分大小写 缺省情况下，MST 域的域名等于交换设备主控板上管理网口的 MAC 地址，即桥 MAC 地址，可用 **undo region-name** 命令恢复交换设备 MST 域名为缺省值

（续表）

<table>
<tr><th>配置任务</th><th>步骤</th><th>命令</th><th colspan="2">说明</th></tr>
<tr><td rowspan="2">配置
并激活
MST 域</td><td rowspan="2">5</td><td>Instance instance-id vlan { vlan-id1 [to vlan-id2] }&<1-10>
例如：[HUAWEI-mst-region] instance 1 vlan 1 to 3</td><td>配置多生成树实例和 VLAN 的映射关系。命令中的参数说明如下。
（1）instance-id：取值范围不同 S 系列交换机有所不同，S2700 系列为 0～16，S3700/9300 系列为 0～48，S5700/6700/7700/9700 为 0～4 094 的整数，取值为 0 表示的是 CIST
（2）vlan-id1：指定要映射的 VLAN 的起始 VLAN ID，取值范围为 1～4 094
（3）vlan-id2：可选参数，指定要映射的 VLAN 的结束 VLAN ID，取值范围为 2～4 094
（4）&<1-10>：表示前面的参数或参数对最多可以重复 10 次
缺省情况下，所有 VLAN 均映射到 CIST，即实例 0 上，可用 undo instance instance-id [vlan { vlan-id1 [to vlan-id2] } &<1-10>]命令删除指定 VLAN 和指定生成树实例的映射关系</td><td rowspan="2">（二选一）配置多生成树实例与 VLAN 的映射关系。由于命令 vlan-mapping modulo modulo 实现的自动分配机制很难刚好满足实际的多生成树实例与 VLAN 的映射关系，建议使用命令 instanceinstance-id vlan { vlan-id1 [to vlan-id2] }&<1-10> 配置多生成树实例和 VLAN 的映射关系</td></tr>
<tr><td>vlan-mapping modulo modulo
例如：[HUAWEI-mst-region] vlan-mapping modulo 2</td><td>配置多生成树实例和 VLAN 按照缺省算法自动分配映射关系。参数 modulo 用来指定映射的模值，取值范围也因不同交换机系列有所不同：S2700 系列为 1～16 的整数，S3700/9300 系列为 1～48 的整数，S5700/6700/ 7700/9700 系列为 1～64 的整数
【说明】本命令可以快速配置 VLAN 映射表，使每个 VLAN 按照配置被映射到不同的生成树实例上。它是指 VLAN ID 减 1 后除以模值 modulo 值的余数再加 1，即 (VLAN ID-1) %modulo+1，然后通过此算法来分配到对应的实例中，即余数加 1 为几就将此 VLAN 分配到实例几中。如模值 modulo 为 16，则 VLAN1 映射到 MSTI1、VLAN2 映射到 MSTI2……VLAN16 映射到 MSTI16、VLAN17 映射到 MSTI1，依此类推
缺省情况下，所有 VLAN 均映射到 CIST，即实例 0 上，可用 undo vlan-mapping modulo 命令将多生成树实例和 VLAN 按照缺省算法自动分配映射关系恢复为缺省情况</td></tr>
</table>

（续表）

配置任务	步骤	命令	说明
配置并激活 MST 域	6	**revision-level** *level* 例如：[HUAWEI-mst-region] **revision-level** 5	配置 MST 域的 MSTP 修订级别，取值范围为 0～65 535 的整数。当设备所在域的 MSTP 修订级别不为 0，则需要执行本操作 缺省情况下，MSTP 域的 MSTP 修订级别为 0
	7	**active region-configuration** 例如：[HUAWEI-mst-region]**active region-configuration**	激活 MST 域的配置，使以上 MST 域名、VLAN 映射表和 MSTP 修订级别配置生效 如果不执行本操作，以上配置的域名、VLAN 映射表和 MSTP 修订级别无法生效。如果在启动 MSTP 特性后又修改了交换设备的 MST 域相关参数，可以通过执行本命令激活 MST 域，使修改后的参数生效 【说明】由于 MST 域相关参数（特别是 VLAN 映射表）的变化会引起 MSTP 重新计算生成树，从而引起网络拓扑振荡。因此，在完成配置 MST 域名、配置多生成树实例与 VLAN 的映射关系和配置 MST 域的 MSTP 修订级别后，建议在 MST 域视图下执行命令 **check region-configuration** 确定未生效的域参数配置是否正确。在确认域参数无误后，再执行本命令激活新的 MST 域配置
（可选）配置根桥和备份根桥	8	**quit** 例如：[HUAWEI-mst-region] **quit**	退出 MST 域视图，返回系统视图
	9	**stp** [**instance** *instance-id*] **root** { **primary** | **secondary** } 例如：[HUAWEI] **stp instance** 1 **root primary**	配置当前设备为指定 MSTI 的根桥或备份根桥。可选参数 *instance-id* 用来指定 MSTI 的编号，如果不指定此可选参数，则将作为 CIST 的根桥或备份根桥设备 配置为根桥后该设备优先级 BID 值自动为 0，配置为备份根桥后该设备优先级 BID 值自动为 4 096，且都不能更改 缺省情况下，交换设备不作为任何生成树的根桥和备份根桥，可用 **undo stp root** 命令取消当前设备作为指定 MSTI 的根桥或备份根桥的资格
（可选）配置交换设备在指定生成树实例中的优先级	10	**stp** [**instance** *instance-id*] **priority** *priority* 例如：[HUAWEI] **stp instance** 1 **priority** 100	配置当前设备在指定 MSTI 中的桥优先级。命令中的参数说明如下。 （1）*instance-id*：可选参数，用来指定 MSTI 的编号，如果不指定此可选参数，则将配置当前设备在 CIST 中的桥优先级 （2）*priority*：指定当前设备的桥优先级，取值范围是 0～61 440，步长为 4 096，即**仅可以配置 16 个优先级取值，如 0、4 096、8 192 等，不能随便设**。优先级值越小，则优先级越高，越能成为根桥或备份根桥 缺省情况下，交换设备的桥优先级值为 32 768，可用 **undo** [**instance** *instance-id*] **stp priority** 命令恢复交换机的桥优先级为缺省值 【注意】如果已执行了上步命令将当前交换机作为根桥或备份根桥，则在需要改变当前设备的优先级时需先执行 **undo stp** [**instance** *instance-id*] **root** 去使能根交换设备或者备份根交换设备功能，然后执行本命令配置新的优先级数值

（续表）

配置任务	步骤	命令	说明
	11	**stp pathcost-standard** { **dot1d-1998** | **dot1t** | **legacy** } 例如：[HUAWEI]**stp pathcost-standard dot1d-1998**	配置端口路径开销缺省值的计算方法。命令中的选项说明如下。 （1）**dot1d-1998**：多选一选项，表示采用 IEEE 802.1D 标准计算方法 （2）**dot1t**：多选一选项，表示采用 IEEE 802.1t 标准计算方法 （3）**legacy**：多选一选项，表示采用华为的私有计算方法 缺省情况下，路径开销缺省值的计算方法为 IEEE 802.1t（**dot1t**）标准方法，可用 **undo stp pathcost-standard** 命令恢复路径开销缺省值采用缺省计算方法。且同一网络内所有交换设备的端口路径开销应使用相同的计算方法
	12	**Interface** *interface-type interface-number* 例如：[HUAWEI] **interface** GigabitEthernet **1/0/0**	进入要参与生成树计算的接口视图
（可选）配置端口在指定生成树实例中的路径开销	13	**stp** [**instance** *instance-id*] **cost** *cost* 例如：[HUAWEI-GigabitEthernet1/0/0] **stp instance** 1 **cost** 200	设置当前端口在指定生成树实例中的路径开销值，用于桥的根端口选举，值越大，优先级越低。命令中的参数说明如下。 （1）*instance-id*：可选参数，指定要设置当前端口路径开销值的所在 MSTI 编号，如果不指定此参数，则为在 CIST 中配置该端口的路径开销值 （2）*cost*：设置当前端口在指定 MSTI 中的路径开销值。取值范围根据所采用的计算方法的不同而不同： ① 使用华为的私有计算方法时参数 *cost* 的取值范围是 1～200 000 ② 使用 IEEE 802.1d 标准方法时参数 *cost* 的取值范围是 1～65 535 ③ 使用 IEEE 802.1t 标准方法时参数 *cost* 的取值范围是 1～200000000 缺省情况下，端口的路径开销值为接口速率对应的路径开销缺省值，可用 **undo stp** [**instance** *instance-id*] **cost** 命令恢复当前接口在指定 MSTI 中的路径开销为缺省值。当采用华为私有计算方法时的缺省值参见表 8-5 【说明】在存在环路的网络环境中，对于链路速率值相对较小的接口，建议将其路径开销值配置相对较大，以使其在生成树算法中被选举成为阻塞端口，阻塞其所在链路，因为开销值越大的接口越将成为阻塞端口
（可选）配置端口优先级	14	**stp** [**instance** *instance-id*] **port priority** *priority* 例如：[HUAWEI-GigabitEthernet1/0/0] **stp port priority** 64	配置端口在指定生成树实例中的优先级，参与指定端口的选举。命令中的参数说明如下。 （1）*instance-id*：可选参数，指定要设置当前端口优先级值的所在 MSTI 编号，如果不指定此参数，则为在 CIST 中配置该端口的优先级值 （2）*priority*：设置当前端口在指定 MSTI 中的优先级，取值范围是 0～240，步长为 16，不能随便设置，且优先级值越小，优先级越高，越能成为指定端口 缺省情况下，端口的优先级取值是 128，可用 **undo stp port priority** 命令恢复当前接口的优先级为缺省值

（续表）

配置任务	步骤	命令	说明
对所有要参与 STP 或者 RSTP 生成树计算的各交换机端口重复以上第 12～第 14 步			
启用 MSTP	15	**quit** 例如：[HUAWEI-GigabitEthernet1/0/0] **quit**	退出接口视图，返回系统视图
	16	**stp enable** 例如：[HUAWEI] **stp enable** 或 [HUAWEI-GigabitEthernet1/0/0] **stp enable**	使能交换机的 MSTP 功能。本命令既可在系统视图下全局启用交换机上各端口的 MSTP 功能，也可在具体接口视图下仅启用对应接口的 MSTP 功能 缺省情况下，S2700/3700/9300 系列交换机中的 STP/RSTPMSTP 功能处于启用状态，可用 **undo stp enable** 命令去使能交换设备或端口上的 STP/RSTP/MSTP 功能。也可用 **stp disable** 命令去使能交换设备或端口上的 MSTP/RSTP 功能（在 S2700/3700 系列交换机中仅可使用此命令去使能 STP/RSTP/MSTP 功能）；在 S5700/6700/7700/9700 系列交换机中的 MSTP 功能处于禁用状态，可用 **stp enable 或者 undo stp disable** 命令使能交换设备或端口上的 STP/RSTP/MSTP 功能
（可选）配置端口的收敛方式	17	**stp converge { fast | normal }** 例如：[HUAWEI] **stp converge fast**	配置 MSTP 多生成树的收敛方式。命令中的选项说明如下。 （1）**fast**：二选一选项，指定采用快速方式，ARP 表将需要更新的表项直接删除 （2）**normal**：二选一选项，指定采用普通模式，仅将 ARP 表中需要更新的表项快速老化 缺省情况下，端口的 STP/RSTP/MSTP 收敛方式为 normal，可用 **undo stp converge** 命令恢复 STP/RSTP/MSTP 收敛方式为缺省值。建议选择 normal 收敛方式。若选择 fast 方式，频繁的 ARP 表项删除可能会导致设备 CPU 占用率高达 100%，报文处理超时导致网络振荡

8.6.3 MSTP多进程基本功能及主要配置任务

在二层单接环和二层双接环混合组网环境下，交换设备同时承载二三层业务。为了实现不同的环完成不同的业务，可在该组网上部署MSTP多进程，实现不同环上的生成树协议进行独立计算，互不影响。但**S2700/3700/5700LI**系列交换机不支持**MSTP**多进程功能 。

在如图8-36所示的网络中，SwitchA、SwitchB和SwitchC之间通过二层链路相连，并且同时启动MSTP协议。环上的CE设备只支持STP/RSTP协议并且存在多个接入环，不同的接入环在SwitchA和SwitchB上通过不同的端口接入。

在图中，SwitchA和 SwitchB 之间的链路也是二层链路，并运行 MSTP 协议，但SwitchA和SwitchB之间的链路是多接入环的共享链路，它与其他接入环链路的不同在于：共享链路上的端口需要参与多个接入环和多个MSTP进程的计算，这样SwitchA 和 SwitchB 之间的 MSTP 协议报文就需要能区分是来自哪个进程的MSTP协议报文。

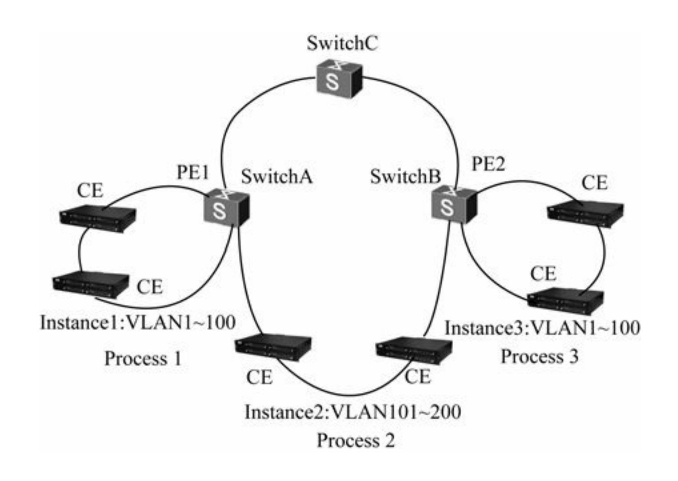


图8-36 MSTP多进程示例

此外，共享链路上的同一个端口同时参与多个 MSTP 进程的计算，多个MSTP进程中都会计算出端口状态，这样端口就可能同时存在多个状态，从而无法决定采用哪个生成树的状态。要注意的是，虽然共享链路上的端口参与了多个 MSTP 进程的状态计算，但是该共享链路只具有MSTP进程0的状态，不会影响其他MSTP进程。

MSTP的多进程功能实现了各个环之间的 MSTP 独立运行，每个 MSTP进程可以管理设备上的部分端口，即设备的二层端口资源被多个MSTP进程分割管理，每个MSTP进程上都运行标准的MSTP协议。但在配置MSTP多进程基本功能之前，需完成MST域配置并激活。

MSTP多进程下的基本功能配置任务如下。总体上与8.6.1节介绍的单进程下MSTP基本功能配置任务存在许多类似之处，只是需要创建多个MSTP进程，然后在对应进程下为对应的 MSTI 配置 MSTP 基本功能，而且必须先完成 **8.6.1** 节介绍的单进程下的**MSTP**工作模式和**MST**域名并激活配置 。

（1）创建MSTP进程。进程的ID是识别MSTP多进程的唯一标识。MSTP设备将端口绑定在进程中，设备将以进程为边界进行MSTP协议计算，不在此进程内的端口将不参与此进程的协议计算。需要在与多个接入环相连接的设备上进行MSTP进程创建配置。

（2）配置端口加入MSTP进程。通过将端口加入MSTP进程中，可使其参与MSTP协议计算。使能MSTP功能的设备与接入环相连的链路叫做接入链路，多个接入环共用的链路叫做共享链路。共享链路上的端口需要参与多个接入环和多个MSTP进程的计算，所以需要加入多个MSTP进程。

（3）（可选）配置根桥和备份根桥。可以通过计算自动确定生成树的根桥，用户也可以手动配置设备为指定MSTP进程下的指定MSTI配置根桥或备份根桥。

（4）（可选）配置交换设备在指定生成树实例中的优先级。可以为各交换设备配置在指定MSTP进程下的指定MSTI中的桥优先级，以便参与在指定MSTI中的根桥选举。在进行根桥的选择时，一般希望选择性能高、网络层次高的交换设备作为根桥。但是，性能高、网络层次高的交换设备其优先级不一定高，因此需要配置优先级以保证该设备成为根桥。

（5）（可选）配置端口在指定生成树实例中的路径开销。同样可以为各运行 MSTP协议的端口配置在指定 MSTP 进程下的指定 MSTI 中的路径开销值，以便参与在指定MSTI中的根端口的选举。

（6）（可选）配置端口在指定生成树实例中的优先级。这里是为各运行 MSTP 协议的端口配置在指定MSTP进程下的指定MSTI中的优先级值，以便参与在指定MSTI中的指定端口的选举。

（7）配置MSTP多进程的TC通告功能。配置MSTP多进程的TC通告功能后，当前MSTP进程在收到TC报文后，能够及时通告给其他指定MSTP进程中的实例，以便 使其及时刷新MAC表项和ARP表项，从而保证用户业务不中断。这是需要在与接入环相连接的设备上进行配置。

（8）启用MSTP。当交换设备配置MSTP多进程基本功能后，必须在指定进程视图下使能设备MSTP功能，该进程的MSTP相关配置才能生效。

8.6.4 配置MSTP多进程基本功能

上节介绍的MSTP多进程基本功能八项配置任务的具体配置步骤如表8-19所示。

表8-19 MSTP基本功能配置步骤

配置任务	步骤	命令	说明
创建 MSTP 进程	1	**system-view** 例如：<HUAWEI> **system-view**	进入系统视图
	2	**stp process** *process-id* 例如：[HUAWEI] **stp process** 10	创建 MSTP 进程并进入该 MSTP 进程的视图。S5700/6700 系列交换机的取值范围为 1～25 的整数；S7700/9300/9700 系列交换机的取值范围 0 为 1～63 的数，最大支持 16 个 MSTP 进程 缺省情况下，整个设备的所有 MSTP 相关配置均属于进程 0，可用 **undo stp process** *process-id* 命令删除一个指定的 MSTP 进程
配置并激活 MST 域	3	**quit** 例如：[HUAWEI-mst-process-10] **quit**	退出 MSTP 进程视图，返回系统视图
	4	**interface** *interface-type interface-number* 例如：[HUAWEI] **interface** gigabitethernet 1/0/1	键入要加入 MSTP 进程的交换机端口，进入接口视图
	5	**stp binding process** *process-id* 例如：[HUAWEI-GigabitEthernet1/0/1] **stp binding process** 10	（二选一）在接入链路上将当前接口加入指定 ID 的 MSTP 进程中。如果加入 MSTP 进程的端口上存在子接口，并且子接口上配置了其他业务，例如 VPLS 业务，此时可以在主接口上使用命令 **stp vpls-subinterface enable**，当主接口在收到 TC 报文后，能够通告其子接口及时刷新 MAC 表项和 ARP 表项，从而保证用户业务不中断。另外还需要在主接口上配置 Root 保护。 一个接入链路所在接口只能加入一个 MSTP 进程，若多次执行本命令配置当前接口加入不同 ID 的 MSTP 进程，以最后一次配置为准 缺省情况下，接口属于 ID 为 0 的 MSTP 进程，可用 **undo stp binding process** *process-id* 命令将当前接口退出指定 ID 的 MSTP 进程中
	6	**stp binding process** *process-id1* [**to** *process-id2*] **link-share** 例如：[HUAWEI-GigabitEthernet1/0/1] **stp binding process** 10 **to** 12	（二选一）在共享链路上将共享链路上的端口加入多个 MSTP 进程。此命令中指定的端口不是设备与接入环相连接的接口，而是配置了 MSTP 多进程的设备之间的共享链路上的接口 对于存在共享链路的进程，必须在多进程视图下使能 **stp enable**。对于以共享链路方式加入进程的端口，必须在端口下使能 **stpenable**

（续表）

配置任务	步骤	命令	说明
（可选）配置根桥和备份根桥	7	**quit** 例如：[HUAWEI-GigabitEthernet1/0/1] **quit**	退出接口视图，返回系统视图
	8	**stp process** *process-id* 例如：[HUAWEI] **stp process** 10	进入已创建，并且配置根桥或备份根桥的 MSTP 进程视图
	9	**stp** [**instance** *instance-id*] **root** {**primary** | **secondary**} 例如：[HUAWEI-mst-process-10] **stp instance** 1 **root primary**	配置当前设备为指定 MSTP 进程下指定 MSTI 的根桥或备份根桥。可选参数 *instance-id* 用来指定 MSTI 的编号，如果不指定此可选参数，则将作为 CIST 的根桥或备份根桥设备 配置为根桥后该设备优先级 BID 值自动为 0，配置为备份根桥后该设备优先级 BID 值自动为 4 096，且都不能更改 缺省情况下，交换设备不作为任何生成树的根桥和备份根桥，可用 **undo stp root** 命令取消当前设备作为指定 MSTI 的根桥或备份根桥的资格
（可选）配置交换设备在指定生成树实例中的优先级	10	**stp** [**instance** *instance-id*] **priority** *priority* 例如：[HUAWEI-mst-process-10]**stp instance** 1 **priority** 8192	配置当前设备在指定 MSTP 进程下指定 MSTI 中的桥优先级。命令中的参数说明如下。 （1）*instance-id*：可选参数，用来指定 MSTI 的编号，如果不指定此可选参数，则将配置当前设备在 CIST 中的桥优先级 （2）*priority*：指定当前设备的桥优先级，取值范围是 0～61 440，步长为 4 096，即**仅可以配置 16 个优先级取值，如 0、4 096、8 192 等，不能随便设**。优先级值越小，则优先级越高，越能成为根桥或备份根桥 缺省情况下，交换设备的桥优先级值为 32 768，可用 **undo** [**instance** *instance-id*] **stp priority** 命令恢复交换机的桥优先级为缺省值 【注意】如果已执行了上步命令将当前交换机作为根桥或备份根桥，则在需要改变当前设备的优先级时需先执行 **undo stp** [**instance** *instance-id*] **root** 去使能根交换设备或者备份根交换设备功能，然后执行本命令配置新的优先级数值
（可选）配置端口在指定生成树实例中的路径开销	11	**quit** 例如：[HUAWEI-mst-process-10] **quit**	退出 MSTP 进程视图，返回系统视图
	12	**stp pathcost-standard** { **dot1d-1998** | **dot1t** | **legacy** } 例如：[HUAWEI]**stp pathcost-standard dot1d-1998**	配置端口路径开销缺省值的计算方法。命令中的选项说明参见表 8-19 中第 11 步的说明
	13	**interface** *interface-type interface-number* 例如：[HUAWEI]**interface** GigabitEthernet 1/0/2	进入要参与生成树计算，配置路径开销的端口的接口视图
	14	**stp binding process** *process-id* 例如：[HUAWEI-GigabitEthernet1/0/2]**stp binding process** 10	将当前端口加入指定 MSTP 进程中

（续表）

配置任务	步骤	命令	说明
（可选）配置端口在指定生成树实例中的路径开销	15	**stp** [**process** *process-id*] [**instance** *instance-id*] **cost** *cost* 例如：[HUAWEI-GigabitEthernet1/0/2]**stp instance** 1 **cost** 200	设置当前端口在指定 MSTP 进程下指定生成树实例中的路径开销值，用于桥的根端口选举，值越大，优先级越低。命令中的参数说明如下。 （1）*instance-id*：可选参数，指定要设置当前端口路径开销值的所在 MSTI 编号，如果不指定此参数，则为在 CIST 中配置该端口的路径开销值 （2）*cost*：设置当前端口在指定 MSTI 中的路径开销值。取值范围根据所采用的计算方法的不同而不同。 ➢ 使用华为的私有计算方法时参数 *cost* 的取值范围是 1～200 000 ➢ 使用 IEEE 802.1d 标准方法时参数 *cost* 的取值范围是 1～65 535 ➢ 使用 IEEE 802.1t 标准方法时参数 *cost* 的取值范围是 1～200 000 000 缺省情况下，端口的路径开销值为接口速率对应的路径开销缺省值，可用 **undo stp** [**instance** *instance-id*] **cost** 命令恢复当前接口在指定 MSTI 中的路径开销为缺省值。当采用华为私有计算方法时缺省值参见表 8-5 【说明】在存在环路的网络环境中，对于链路速率值相对较小的接口，建议将其路径开销值配置相对较大，以使其在生成树算法中被选举成为阻塞端口，阻塞其所在链路，因为开销值越大的接口越将成为阻塞端口
（可选）配置端口优先级	16	**stp** [**process** *process-id*] [**instance** *instance-id*] **port priority** *priority* 例如：[HUAWEI-GigabitEthernet1/0/2] **stp process** 10 **instance** 1 **port priority** 64	配置当前端口在指定 MSTP 进程下指定生成树实例中的优先级，参与指定端口的选举。命令中的参数说明如下。 （1）*instance-id*：可选参数，指定要设置当前端口优先级值的所在 MSTI 编号，如果不指定此参数，则为在 CIST 中配置该端口的优先级值 （2）*priority*：设置当前端口在指定 MSTI 中的优先级，取值范围是 0～240，步长为 16，不能随便设置，且优先级值越小，优先级越高，越能成为指定端口 缺省情况下，端口的优先级取值是 128，可用 **undo stp port priority** 命令恢复当前接口的优先级为缺省值
对所有要参与 STP 或者 RSTP 生成树计算的各交换机端口重复以上第 12～第 16 步			
配置 MSTP 多进程的 TC 通告功能	17	**quit** 例如：[HUAWEI-GigabitEthernet1/0/2] **quit**	退出接口视图，返回系统视图
	18	**stp process** *process-id* 例如：[HUAWEI] **stp process** 10	进入已创建，要配置 TC 通告功能的 MSTP 进程的 MSTP 进程视图
	19	**stp tc-notify process 0** 例如：[HUAWEI-mst-process-10] **stp tc-notify process 0**	使能当前 MSTP 进程的 TC 通告功能。执行本命令后，当前 MSTP 进程在收到 TC 报文后能够及时通告给 MSTP 进程 0 中的实例，以便使其及时刷新 MAC 表项和 ARP 表项，从而保证用户业务不中断

（续表）

配置任务	步骤	命令	说明
启用 MSTP	20	**stp enable** 例如：[HUAWEI-mst-process-10] **stp enable**	在对应 MSTP 进程下使能交换机的 MSTP 功能 缺省情况下，S2700/3700/9300 系列交换机中的 STP/RSTPMSTP 功能处于启用状态，可用 **undo stp enable** 命令去使能交换设备上的 STP/RSTP/MSTP 功能。也可用 **stp disable** 命令去使能交换设备上的 MSTP/RSTP 功能（在 S2700/3700 系列交换机中仅可使用此命令去使能 STP/RSTP/MSTP 功能）；在 S5700/6700/7700/9700 系列交换机中的 MSTP 功能处于禁用状态，可用 **stp enable 或者 undo stp disable** 命令使能交换设备上的 STP/RSTP/MSTP 功能
（可选）配置端口的收敛方式	21	**quit** 例如：[HUAWEI-mst-process-10] **quit**	退出 MSTP 进程视图，返回系统视图
	22	**stp converge { fast | normal }** 例如：[HUAWEI] **stp converge fast**	配置 MSTP 多生成树的收敛方式。命令中的选项说明如下。 （1）**fast**：二选一选项，指定采用快速方式，ARP 表将需要更新的表项直接删除 （2）**normal**：二选一选项，指定采用普通模式，仅将 ARP 表中需要更新的表项快速老化 缺省情况下，端口的 STP/RSTP/MSTP 收敛方式为 normal，可用 **undo stp converge** 命令恢复 STP/RSTP/MSTP 收敛方式为缺省值。建议选择 normal 收敛方式。若选择 fast 方式，频繁的 ARP 表项删除可能会导致设备 CPU 占用率高达 100%，报文处理超时导致网络振荡

8.6.5 配置影响**MSTP**拓扑收敛的参数

配置合适的影响MSTP拓扑收敛的参数来实现最快速度的拓扑收敛。在配置MSTP影响拓扑收敛的参数之前，需完成上节介绍的MSTP基本功能配置。

整个影响MSTP拓扑收敛的参数配置与RSTP中的参数配置极其相似，主要也是网络直径、超时时间、定时器、影响带宽的最大连接数、端口的链路类型、端口的最大发送速率、执行MCheck操作、边缘端口和BPDU报文过滤功能等，与**RSTP**中的配置不同的只是原来在系统视图中的配置现在需要在对应的**MSTP**进程视图下进行配置，除了在全局模式下配置边缘端口和**BPDU**报文过滤功能，参见**8.4.5**节表**8-9** 。

另外，在MSTP中还可配置MST域内生成树的最大跳数。在MST BPDU中包含一个记录该BPDU剩余生存跳数（CIST Remaining Hops）字段，参见 8.5.6节表8-15。MST域内生成树的最大跳数决定了生成树的网络规模大小，从而控制生成树的网络规模。

剩余生存跳数的计算方法如下。

（1）根桥设备发送的BPDU的剩余生存跳数为MST域的最大跳数。

（2）非根桥设备发送的BPDU的剩余生存跳数为MST域的最大跳数减去本桥设备距根桥设备的跳数。

（3）如果交换设备收到的BPDU中携带的剩余生存跳数为0，则交换设备将该BPDU丢弃。配置MST域的最大跳数的方法是在对应MSTP进程视图或系统视图下（在**S2700/3700**系列交换机中仅可在系统视图下配置 ）使用stp max-hopshop命令进行配置，取值范围为1～40 的整数。缺省情况下，MST域内生成树的最大跳数为20，可用undo stp max-hops命令恢复MST域内生成树的最大跳数为缺省值。但本配置仅需要在ID非0的MSTP进程中进行。

【示例**1** 】 配置MST域内所有生成树的最大跳数为35。

<HUAWEI>system-view

[HUAWEI] stp max-hops 35

【示例 **2** 】配置MST域内MSTP进程10中的所有生成树的最大跳数为35。

<HUAWEI>system-view

[HUAWEI] stp process 10

[HUAWEI-mst-process-10] stp max-hops 35

8.6.6 配置MSTP保护功能

MSTP也支持RSTP所有的保护功能，包括BPDU保护功能、防TC-BPDU报文攻击保护功能、Root 保护功能和环路保护功能，参见 8.4.6 节介绍。另外，MSTP 还提供了特有的共享链路保护功能。“共享链路保护功能”用在交换设备双归属接入网络的场景中。当共享链路出现故障时，通过共享链路保护功能，使本设备的工作模式强制转换为RSTP，配合使用根保护功能，可以避免网络环路。用户可根据实际环境任选其中一个或多个保护功能配置。具体配置步骤如表8-20所示。

表8-20 RSTP保护功能的配置步骤

配置任务	步骤	命令	说明
配置 BPDU 保护功能	1	**system-view** 例如：<HUAWEI> **system-view**	进入系统视图
	2	**stp process** *process-id* 例如：[HUAWEI] **stp process** 10	进入要配置 BPDU 保护功能的 MSTP 进程的 STP 进程视图
	3	**stp bpdu-protection** 例如：[HUAWEI-mst-process-10] **stp bpdu-protection**	配置边缘端口的 BPDU 保护功能，其他说明参见 8.4.6 节表 8-11 的第 2 步
配置 TC-BPDU 报文攻击保护功能	4	**stp tc-protection** 例如：[HUAWEI-mst-process-10] **stp tc-protection**	使能交换设备在当前 MSTP 进程下对 TC 类型 BPDU 报文的保护功能，其他说明参见 8.4.6 节表 8-11 的第 3 步
	5	**stp tc-protection threshold** *threshold* 例如：[HUAWEI-mst-process-10] **stp tc-protection threshold** 10	配置交换设备在当前MSTP进程下在收到TC类型 BPDU 报文后，单位时间内处理 TC 类型 BPDU 报文，并立即刷新转发表项的阈值，其他说明参见 8.4.6 节表 8-11 的第 4 步
	6	**quit** 例如：[HUAWEI-mst-process-10] **quit**	退出 MSTP 进程视图，返回系统视图
配置端口的 Root 保护功能	7	**interface** *interface-type interface-number* 例如：[HUAWEI]**interface** gigabitethernet 0/0/1	进入指定端口的进口视图
	8	**stp binding process** *process-id* 例如：[HUAWEI-GigabitEthernet0/0/1] **stp binding process** 10	将端口绑定到指定的 MSTP 进程，步骤仅在需要把接口绑定到 ID 非 0 进程时配置。当接口属于 ID 为 0 的进程，可跳过本步，直接进入下一步
	9	**stp root-protection** 例如：[HUAWEI-GigabitEthernet0/0/1] **stp root-protection**	配置以上指定端口（**只能在指定端口下配置**）的 Root 保护功能，其他说明参见 8.4.6 节表 8-11 的第 6 步

（续表）

配置任务	步骤	命令	说明
配置端口的环路保护功能	10	**quit** 例如：[HUAWEI-GigabitEthernet0/0/1] **quit**	退出以上指定端口的接口视图，返回系统视图
	11	**interface** *interface-type interface-number* 例如：[HUAWEI]**interface** gigabitethernet 0/0/2	进入根端口或者 Alternate 端口的接口视图
	12	**stp loop-protection** 例如：[HUAWEI-GigabitEthernet0/0/2]**stp loop-protection**	配置交换设备根端口或 Alternate 端口的环路保护功能，**不能在指定端口下配置**，其他说明参见 8.4.6 节表 8-11 的第 9 步
配置共享链路保护功能	13	**quit** 例如：[HUAWEI-GigabitEthernet0/0/2] **quit**	退出接口视图，返回系统视图
	14	**stp process** *process-id* 例如：[HUAWEI] **stp process** 10	进入要配置 BPDU 保护功能的 MSTP 进程的 STP 进程视图
	15	**stp link-share-protection** 例如：[HUAWEI-mst-process-10] **stp link-share-protection**	使能共享链路保护功能。本步骤仅需要在 ID 非 0 的 MSTP 进程中配置系统参数时执行因为 S2700/3700/5700LI 系列交换机不支持 MSTP 多进程，所以不支持本项配置

8.6.7 配置MSTP支持和其他厂商设备互通的参数

为了实现与其他厂商设备的互通，需要在华为公司运行MSTP的设备上配置一些参数，以确保通信畅通。包括以下几项配置任务。

（1）配置端口Proposal/Agreement机制的迁移方式。与RSTP一样，在华为S系列交换机的MSTP协议中也支持增强模式和普通模式两种端口Proposal/Agreement机制的迁移方式，具体参见8.4.7节介绍。

（2）配置端口收发 MSTP 协议的报文格式。MSTP 协议报文存在两种格式，一种为dot1s，即IEEE802.1s规定的报文格式，另一种是华为私有legacy报文格式。配置时，可以指定报文的格式，也可以配置MSTP协议报文格式自适应的功能，即根据收到的MSTP协议报文格式自动切换接口支持的MSTP协议报文格式，使报文格式与对端匹配。

（3）使能摘要监听功能。当华为设备与其他厂商的设备互连时，在域名、修订级别、VLAN实例映射表全都一致的情况下，由于双方BPDU报文密钥不一致，会导致两台设备不能正常互通，在这种情况下，需要在交换设备上使能摘要监听功能。可在MST域中的交换设备上通过配置实现华为设备的BPDU报文密钥与其他厂商设备的BPDU报文密钥一致。

以上三项配置任务的具体配置步骤如表8-21所示。

表8-21 支持与其他厂商设备互通的参数配置步骤

配置任务	步骤	命令	说明
配置端口P/A 机制的迁移方式	1	**system-view** 例如：<HUAWEI> **system-view**	进入系统视图
	2	**interface** *interface-type interface-number* 例如：[HUAWEI] **interface** gigabitethernet 0/0/1	键入与其他厂商设备直接相连的端口，进入接口视图

（续表）

配置任务	步骤	命令	说明
配置端口P/A 机制的迁移方式	3	**stp no-agreement-check** 例如：[HUAWEI-GigabitEthernet0/0/1]**stp no-agreement-check**	配置端口使用普通的快速迁移方式，缺省情况下，端口使用增强的快速迁移机制，可用 **undo stp no-agreement-check** 命令配置当前接口使用增强的快速迁移机制
配置端口收发 MSTP 协议的报文格式	4	**stp compliance** { **auto** | **dot1s** | **legacy** } 例如：[HUAWEI-GigabitEthernet0/0/1]**stp compliance auto**	配置端口协议报文格式。缺省情况下，MSTP 报文收发格式为 auto 模式。如果直连的接口上一端配置 dot1s，而另一端配置 legacy，是不能协商成功的
使能摘要监听功能	5	**stp config-digest-snoop** 例如：[HUAWEI-GigabitEthernet0/0/1]**stp config-digest-snoop**	使能端口上配置摘要监听的功能

8.6.8 MSTP功能配置示例

本示例拓扑结构如图8-37所示，SwitchA、SwitchB、SwitchC和SwitchD都运行MSTP。它们彼此相连形成了一个环网，因为在SwitchA与SwitchB之间，以及SwitchC与SwitchD之间都存在冗余链路（本来这些链路都是可以不要的）。为实现 VLAN2～VLAN10 和VLAN11～VLAN20的流量负载分担，采用MSTP协议配置了两个MSTI，即MSTI1和MSTI2。

1．配置思路分析

（1）在四台交换机创建一个相同的MST域，然后在这个MST域中创建两个MSTI （MSTI1和MSTI2），它们的生成树拓扑参见图8-37。把ID号为2～20的VLAN映射到MSTI1中，把ID号为11～20的VLAN映射到MSTI2中。

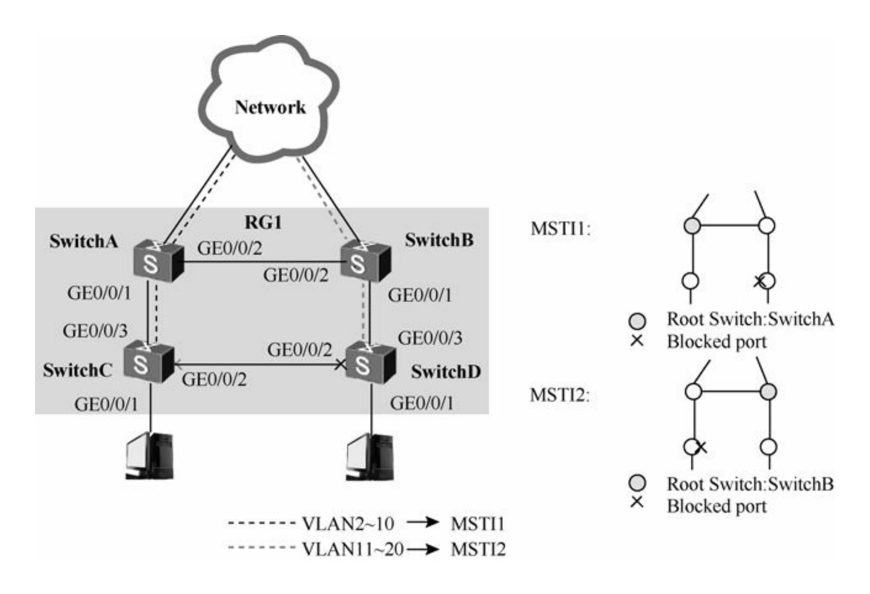


图8-37 MSTP配置示例

（2）为了实现两个MSTI无二层环路，在MSTI1中阻塞了SwitchD上的GE0/0/2端口，在MSTI2中阻塞了SwitchC上的GE0/0/2端口。

（3）配置MSTI的根桥为SwitchA，MSTI2的根桥为SwitchB，这样就实现了MSTI1中的VLAN2～VLAN10和MSTI2中的VLAN11～VLAN20的流量通过上行两条链路进行负载分担。

（4）最后在这台交换机上启用MSTP协议，使以上配置生效。

（5）为了确保两个MSTI中的根桥不会发生变化，分别在SwitchA和SwitchB两指定端口上配置根保护功能。

（6）在各交换机上创建ID号为2～20的共19个VLAN，配置各链路间端口的类型，并允许对应的VLAN 通过。之所以要把 **VLAN** 的创建与配置放在最后，就是为了预防环路的发生，因为如果在启用**MSTP**协议前创建了这些**VLAN**，肯定会发生二层环路的，也起不到负载分担的目的 。

2．具体配置步骤

根据以上配置思路，下面具体介绍它们的配置步骤。

（1）在4台交换机上分别创建一个相同的MST域（域名假设为RG1）、两个多生树实例MSTI1和MSTI2，然后创建ID为2～10的VLAN映射到MSTI1的映射，创建ID为11～20的VLAN映射到MSTI2的映射。并激活MST域配置。

SwitchA上的**MST**域配置：

```
<HUAWEI>system-view
[HUAWEI] sysname SwitchA
[SwitchA] stp region-configuration
[SwitchA-mst-region] region-name RG1
[SwitchA-mst-region] instance 1 vlan 2 to 10
[SwitchA-mst-region] instance 2 vlan 11 to 20
[SwitchA-mst-region] active region-configuration
[SwitchA-mst-region] quit
```

SwitchB上的**MST**域配置：

```
<HUAWEI>system-view
[HUAWEI] sysname SwitchB
[SwitchB] stp region-configuration
[SwitchB-mst-region] region-name RG1
[SwitchB-mst-region] instance1 vlan 2 to 10
[SwitchB-mst-region] instance2 vlan 11 to 20
[SwitchB-mst-region] active region-configuration
[SwitchB-mst-region] quit
```

SwitchC上的**MST**域配置：

```
<HUAWEI>system-view
[HUAWEI] sysname SwitchC
[SwitchC] stp region-configuration
[SwitchC-mst-region] region-name RG1
[SwitchC-mst-region] instance 1 vlan 2 to 10
[SwitchC-mst-region] instance2 vlan 11 to 20
[SwitchC-mst-region] active region-configuration
[SwitchC-mst-region] quit
```

SwitchD上的**MST**域配置：

```
<HUAWEI>system-view
[HUAWEI] sysname SwitchD
[SwitchD] stp region-configuration
[SwitchD-mst-region] region-name RG1
[SwitchD-mst-region] instance 1 vlan 2 to 10
[SwitchD-mst-region] instance2 vlan 11 to 20
[SwitchD-mst-region] active region-configuration
[SwitchD-mst-region] quit
```

（2）配置MSTI1与MSTI2的根桥与备份根桥。

配置MSTI1的根桥与备份根桥

```
[SwitchA] stp instance1 root primary    #---配置SwitchA为MSTI1的根桥
[SwitchB] stp instance1 root secondary    #---配置SwitchB为MSTI1的备份根桥
[SwitchB] stp instance2 root primary
[SwitchA] stp instance 2 root secondary
```

（3）配置MSTI1和MSTI2中要被阻塞的端口，以便消除二层环路。

因为本示例中其他端口都是采用对应类型端口的缺省路径开销值，所以要阻塞某端口时只需要把它们的路径开销值配置为大于缺省值即可。路径开销值越大，成为根端口的可能性就越小。

端口路径开销值取值范围由路径开销计算方法决定，这里选择使用华为私有计算方法为例，配置实例MSTI1 和 MSTI2 中将被阻塞的端口（分别为 SwitchD 中的 GE0/0/2 和SwitchC中的GE0/0/2端口）的路径开销值为 20 000（千兆以太网端口路径开销值的缺省值为2）。要求同一网络内所有交换设备的端口路径开销应使用相同的计算方法。下面依次是SwitchA、SwitchB、SwitchC和SwitchD这4台交换机上端口路径开销的相关配置。

[SwitchA] stp pathcost-standard legacy　#---配置采用华为的私有端口路径开销计算方法

[SwitchB] stp pathcost-standard legacy

[SwitchC] stp pathcost-standard legacy

[SwitchC] interfacegigabitethernet 0/0/2

[SwitchC-GigabitEthernet0/0/2] stp instance 2 cost 20000　#---将端口GE0/0/2在实例MSTI2中的路径开销值配置为20000

[SwitchC-GigabitEthernet0/0/2] quit

[SwitchD] stp pathcost-standard legacy

[SwitchD] interfacegigabitethernet 0/0/2

[SwitchD-GigabitEthernet0/0/2] stp instance 1 cost 20000

[SwitchD-GigabitEthernet0/0/2] quit

（4）在4台交换机上全局使能MSTP，使以上MSTP配置生效，消除二层环路。

[SwitchA] stp enable

[SwitchB] stp enable

[SwitchC] stp enable

[SwitchD] stp enable

（5）将与终端PC相连的端口去使能MSTP。

[SwitchC] interface gigabitethernet 0/0/1

[SwitchC-GigabitEthernet0/0/1] stp disable

[SwitchC-GigabitEthernet0/0/1] quit

[SwitchD] interfacegigabitethernet 0/0/1

[SwitchD-GigabitEthernet0/0/1] stp disable

[SwitchD-GigabitEthernet0/0/1] quit

（6）在两实例的根桥设备的指定端口上配置根保护功能。

[SwitchA] interfacegigabitethernet 0/0/1

[SwitchA-GigabitEthernet0/0/1] stp root-protection

[SwitchA-GigabitEthernet0/0/1] quit

[SwitchB] interfacegigabitethernet 0/0/1

[SwitchB-GigabitEthernet0/0/1] stp root-protection

[SwitchB-GigabitEthernet0/0/1] quit

（7）最后在各交换机上创建ID号为2～20的共19个VLAN，然后把4台交换机间的直连链路的端口配置为Trunk类型，并允许这19个VLAN通过。把连接PC的链路端口设置为Access类型，加入对应的VLAN。有关VLAN的具体创建和配置方法参见本书第6章。

SwitchA上的配置：

[SwitchA] vlan batch 2 to 20

[SwitchA] interfacegigabitethernet 0/0/1

[SwitchA-GigabitEthernet0/0/1] port link-type trunk

[SwitchA-GigabitEthernet0/0/1] port trunk allow-pass vlan 2 to 20

[SwitchA-GigabitEthernet0/0/1] quit

[SwitchA] interface gigabitethernet 0/0/2

[SwitchA-GigabitEthernet0/0/2] port link-type trunk
[SwitchA-GigabitEthernet0/0/2] port trunk allow-pass vlan 2 to 20
[SwitchA-GigabitEthernet0/0/2] quit

SwitchB上的配置：

[SwitchB] vlan batch 2 to 20
[SwitchB] interface gigabitethernet 0/0/1
[SwitchB-GigabitEthernet0/0/1] port link-type trunk
[SwitchB-GigabitEthernet0/0/1] port trunk allow-pass vlan 2 to 20
[SwitchB-GigabitEthernet0/0/1] quit
[SwitchB] interface gigabitethernet 0/0/2
[SwitchB-GigabitEthernet0/0/2] port link-type trunk
[SwitchB-GigabitEthernet0/0/2] port trunk allow-pass vlan 2 to 20
[SwitchB-GigabitEthernet0/0/2] quit

SwitchC上的配置：

[SwitchC] vlan batch 2 to 20
[SwitchC] interface gigabitethernet 0/0/1
[SwitchC-GigabitEthernet0/0/1] port link-type access
[SwitchC-GigabitEthernet0/0/1] port default vlan 2
[SwitchC-GigabitEthernet0/0/1] quit
[SwitchC] interface gigabitethernet 0/0/2
[SwitchC-GigabitEthernet0/0/2] port link-type trunk
[SwitchC-GigabitEthernet0/0/2] port trunk allow-pass vlan 2 to 20
[SwitchC-GigabitEthernet0/0/2] quit
[SwitchC] interface gigabitethernet 0/0/3
[SwitchC-GigabitEthernet0/0/3] port link-type trunk
[SwitchC-GigabitEthernet0/0/3] port trunk allow-pass vlan 2 to 20
[SwitchC-GigabitEthernet0/0/3] quit

SwitchD上的配置：

[SwitchD] vlan batch 2 to 20
[SwitchD] interfacegigabitethernet 0/0/1
[SwitchD-GigabitEthernet0/0/1] port link-type access
[SwitchD-GigabitEthernet0/0/1] port default vlan 11
[SwitchD-GigabitEthernet0/0/1] quit
[SwitchD] interface gigabitethernet 0/0/2
[SwitchD-GigabitEthernet0/0/2] port link-type trunk
[SwitchD-GigabitEthernet0/0/2] port trunk allow-pass vlan 2 to 20
[SwitchD-GigabitEthernet0/0/2] quit
[SwitchD] interface gigabitethernet 0/0/3
[SwitchD-GigabitEthernet0/0/3] port link-type trunk
[SwitchD-GigabitEthernet0/0/3] port trunk allow-pass vlan 2 to 20

[SwitchD-GigabitEthernet0/0/3] quit

经过以上配置，在网络计算稳定后可使用以下 display 命令验证配置结果。如在SwitchA上执行display stp brief命令可查看端口状态和端口的保护类型，结果如下。从中可以看到，在MSTI1中，由于SwitchA是根桥，其GE0/0/2和GE0/0/1端口成为指定端口（其中在GE0/0/1端口上配置了根保护）；在MSTI2中，SwitchA为非根桥，其GE0/0/1端口成为指定端口，端口GE0/0/2端口成为根端口。符合本示例中两MSTI生成树拓扑要求。

```
[SwitchA] display stp brief
MSTID  Port                  Role   STP State     Protection
  0    GigabitEthernet0/0/1  DESI   FORWARDING    ROOT
  0    GigabitEthernet0/0/2  DESI   FORWARDING    NONE
  1    GigabitEthernet0/0/1  DESI   FORWARDING    ROOT
  1    GigabitEthernet0/0/2  DESI   FORWARDING    NONE
  2    GigabitEthernet0/0/1  DESI   FORWARDING    ROOT
  2    GigabitEthernet0/0/2  ROOT   FORWARDING    NONE
```

在SwitchB上执行display stp brief命令，结果如下。从中可以看出，在MSTI2中，由于SwitchB是根桥，其GE0/0/1和GE0/0/2端口为指定端口（其中在GE0/0/1端口上配置了根保护）；在MSTI1中，SwitchB为非根桥，其GE0/0/1端口成为指定端口， GE0/0/2端口成为根端口，符合本示例中两MSTI生成树拓扑要求。

```
[SwitchB] display stp brief
MSTID  Port                  Role   STP State     Protection
  0    GigabitEthernet0/0/1  DESI   FORWARDING    ROOT
  0    GigabitEthernet0/0/2  ROOT   FORWARDING    NONE
  1    GigabitEthernet0/0/1  DESI   FORWARDING    ROOT
  1    GigabitEthernet0/0/2  ROOT   FORWARDING    NONE
  2    GigabitEthernet0/0/1  DESI   FORWARDING    ROOT
  2    GigabitEthernet0/0/2  DESI   FORWARDING    NONE
```

在SwitchC上执行display stp interface brief命令，结果如下。从中可以看出，SwitchC的GE0/0/3端口在MSTI1和MSTI2中均为根端口， GE0/0/2端口在MSTI2中被阻塞，在MSTI1中被计算为指定端口，也符合本示例中两MSTI生成树拓扑要求。

```
[SwitchC] display stp interfacegigabitethernet 0/0/3 brief
MSTID  Port                  Role   STP State     Protection
  0    GigabitEthernet0/0/3  ROOT   FORWARDING    NONE
  1    GigabitEthernet0/0/3  ROOT   FORWARDING    NONE
  2    GigabitEthernet0/0/3  ROOT   FORWARDING    NONE
[SwitchC] display stp interfacegigabitethernet0/0/2 brief
MSTID  Port                  Role   STP State     Protection
  0    GigabitEthernet0/0/2  DESI   FORWARDING    NONE
  1    GigabitEthernet0/0/2  DESI   FORWARDING    NONE
  2    GigabitEthernet0/0/2  ALTE   DISCARDING    NONE
```

在SwitchD上执行display stp interface brief命令，结果如下。从中可以看出，SwitchD的GE0/0/3端口在MSTI1和MSTI2中均为根端口， GE0/0/2端口在MSTI1中被阻塞，在MSTI2中被计算为指定端口。

[SwitchD] display stp interfacegigabitethernet 0/0/3 brief

MSTID	Port	Role	STP State	Protection
0	GigabitEthernet0/0/3	ALTE	DISCARDING	NONE
1	GigabitEthernet0/0/3	ROOT	FORWARDING	NONE
2	GigabitEthernet0/0/3	ROOT	FORWARDING	NONE

[SwitchD] display stp interfacegigabitethernet0/0/2 brief

MSTID	Port	Role	STP State	Protection
0	GigabitEthernet0/0/2	ROOT	FORWARDING	NONE
1	GigabitEthernet0/0/2	ALTE	DISCARDING	NONE
2	GigabitEthernet0/0/2	DESI	FORWARDING	NONE

8.7 STP/RSTP/MSTP配置管理

在配置好STP、RSTP或者MSTP功能后，可在任意视图下通过以下display命令查看 STP/RSTP/MSTP 相关配置信息，也可查看拓扑变化相关的统计信息，如果设备拓扑变化次数递增，则可以确定网络存在振荡。还可在用户视图下通过以下 reset 命令清除STP、RSTP或MSTP的统计信息。因为STP、RSTP和MSTP都是生成树协议，且使用的管理命令都一样，只是在输出信息中仅显示设备上所运行的对应生成树协议信息。

（1）使用 display stp [process process-id] [instance instance-id] [interface interface-type interface-number | slot slot-id] [brief] 命令可查看生成树的状态和统计信息。

（2）使用display stp [process process-id] [instance instance-id] topology-change命令可查看STP/RSTP、MSTP拓扑变化相关的统计信息。

（3）使用display stp [process process-id] [instance instance-id] [interface interface- type interface-number | slot slot-id] tc-bpdu statistics命令可查看端口TC/TCN报文收发计数。

（4）使用display stp [process process-id] global命令可查看生成树协议的全局概要信息。当设备的端口较多时，使用前面介绍的 display stp命令查看生成树协议时会显示大量的信息，很难定位比较关心的关键信息，查看全局信息不方便；使用前面介绍的display stp brief命令时只能查看端口的生成树状态信息，无法查看全局信息。通过执行本命令可非常方便地查看生成树协议的全局概要信息。命令中的参数process-id用来指定要查看生成树协议的全局概要信息的MSTP进程号，该命令部分机型不支持，具体可以参考对应的产品手册 。

（5）使用display stp [process process-id] region-configuration [digest] 命令可查看交换设备上当前生效的 MST 域配置信息，包括域名、域的修订级别、VLAN 与生成树实例的映射关系以及配置的摘要。本命令仅适用于**MSTP**协议 。

（6）使用display stp [process process-id] bridge { root | local }命令可查看桥的生成树状态详细信息。可选参数process-id用来指定要查看桥的生成树状态详细信息的MSTP进程ID，如果不指定此可选参数，则显示MSTP进程0中的相关信息。该命令部分机型不支持，具体可以参考对应的产品手册。

（7）使用 reset stp [interface interface-type interface-number] statistics命令可清除生成树的统计信息。在某些情况下，需要统计一定时间内某接口的流量信息，这时必须在统计开始前清除该接口原有的统计信息，使接口重新进行统计。当指定可选参数interface-type interface-number时仅清除对应端口上的生成树的统计信息，否则清除当前设备上所有端口生成树协议的统计信息。

（8）使用 **reset stp error packet statistics** 命令可清除生成树协议的错误报文计数。如果需要观察从当前时间开始的某段时间内生成树协议的错误报文计数，可先使用本命令清除历史统计信息。执行本命令后，所有生成树协议的错误报文计数将清零。

第9章　ACL配置与管理

ACL是由一个或多个用于报文过滤的规则组成的规则集合，通过在不同功能上的应用可达到不同的应用效果，如在本书第3章介绍的各种用户登录控制和网络访问控制，以及QoS流策略、路由信息过滤、策略路由等诸多方面。

在华为S系列交换机中，根据ACL所过滤的报文类型和功能的不同又分为多种ACL类型，如基本ACL、高级ACL、基本ACL6、高级ACL6、二层ACL、用户自定义ACL等。这些ACL所支持的过滤参数各有不同，其中高级ACL（包括高级ACL6）所支持的过滤参数最多，当然应用也最灵活，可以选择不同的参数组合进行报文过滤，以实现特定的目的。

本章主要介绍华为S系列交换机上的基本ACL、高级ACL、二层ACL、用户自定义ACL（因篇幅原因，本章不介绍基于IPv6协议的基本ACL6和高级ACL6），以及这些ACL在简化QoS流策略方面应用的配置与管理方法。另外，在本章最后还将介绍一种特殊的动态ACL技术——自反ACL，以及其配置与管理方法。在交换机端口上使能了自反ACL功能后，交换机可以根据始发流量中应用的TCP、UDP、ICMP这三种协议类型的高级ACL（包括高级ACL6）自动生成一个允许响应报文通过所需的自反ACL。自反ACL规则中的源IP地址与目的IP地址、源端口与目的端口均与始发流量中的源IP地址与目的IP地址、源端口与目的端口对应互换，主要用于内、外网间的访问，其目的是在允许响应流量通过的同时，禁止非法的外网用户主动发起的对内网资源的访问。

9.1 ACL基础

ACL（Access Control List，访问控制列表）是一组报文过滤规则的集合，以允许或阻止符合特定条件的报文通过。ACL的应用非常广泛且非常灵活，在许多领域都可见到它的身影，如内/外网用户的网络访问控制、路由信息过滤、QoS流策略、IPSec报文加密过滤、策略路由等。当ACL被其他功能引用时，根据设备在实现该功能时的处理方式（硬件处理或者软件处理），ACL可以被分为基于硬件的应用和基于软件的应用。

9.1.1 ACL的分类及主要应用

ACL的类型根据不同的划分规则可以有不同的分类。在华为S系列交换机中，按照创建ACL时的命名方式分数字型ACL和命名型ACL：创建ACL时如果仅指定了一个编号，则所创建的是数字型ACL；创建ACL时如果指定了一个名称，则所创建的是命名型ACL。按照ACL功能的不同，又可把ACL分为基于接口的ACL、基本ACL、基本ACL6、高级ACL、高级ACL6、二层ACL和用户自定义ACL这几类。它们的主要区别就是所支持的过滤条件的不同，具体说明如表 9-1 所示。本章仅介绍目前主流应用的基本**ACL**、高级**ACL**、二层**ACL**和用户自定义**ACL**这四类**ACL**的配置与管理方法。在华为**S**系列交换机中，没有与规则匹配的报文将丢弃，这一点要特别注意 。

表9-1 按功能划分的ACL类型

<table>
<tr><th>ACL 类型</th><th>编号范围</th><th>适用的 IP 版本</th><th colspan="2">规则过滤条件</th></tr>
<tr><td>基于接口的 ACL</td><td>1 000～19 999</td><td>IPv4&IPv6</td><td colspan="2">根据 IP 报文的入接口定义规则，实现对报文的匹配过滤</td></tr>
<tr><td>基本 ACL</td><td rowspan="2">2 000～2 999</td><td>IPv4</td><td>可使用 IPv4 报文的源 IP 地址、分片标记和时间段信息来定义规则</td><td rowspan="2">过滤规则较简单</td></tr>
<tr><td>基本 ACL6</td><td>IPv6</td><td>可使用 IPv6 报文的源 IP 地址、分片标记和时间段信息来定义规则</td></tr>
<tr><td>高级 ACL</td><td rowspan="2">3 000～3 999</td><td>IPv4</td><td>既可使用 IPv4 报文的源 IP 地址，也可使用目的地址、IP 优先级、ToS、DSCP、IP 承载的协议类型、ICMP 类型、TCP 源端口/目的端口、UDP 源端口/目的端口号等来定义规则</td><td rowspan="2">过滤规则很复杂，但应用最灵活、最广泛</td></tr>
<tr><td>高级 ACL6</td><td>IPv6</td><td>既可以使用 IPv6 报文数据包的源地址，也可以使用目的地址、IP 承载的协议类型、TCP 的源端口/目的端口、ICMPv6 协议的类型、ICMPv6 Code 等来定义规则</td></tr>
<tr><td>二层 ACL</td><td>4 000～4 999</td><td>IPv4&IPv6</td><td colspan="2">可根据 IP 报文的以太网帧头信息来定义规则，如根据源 MAC 地址、目的 MAC 地址、以太帧协议类型等。过滤规则较简单</td></tr>
<tr><td>用户自定义 ACL</td><td>5 000～5 999</td><td>IPv4&IPv6</td><td colspan="2">可根据偏移位置和偏移量从 IP 报文中提取出一段内容进行匹配过滤，应用于一些特定的环境和需求下，如要过滤网络中传输的包含某段内容信息的数据报文</td></tr>
</table>

说明

基本ACL和高级ACL只对IPv4报文生效，但基本ACL和基本ACL6、高级ACL和高级ACL6对应的编号可以相同，二者互不影响。

ACL可以应用在许多方面，而且应用非常频繁。表9-2所示的是ACL比较典型的应用场景。

表9-2 ACL的典型应用场景

ACL 类型	ACL 应用类型	应用场景
基本 ACL 和基本 ACL6	设备管理中的应用	设备管理中 ACL 的应用场景主要包括以下几种。 （1）当设备作为 FTP、TFTP 服务器时，为提高其安全性，可以通过配置基本 ACL 和基本 ACL6 实现只允许满足过滤条件的客户端访问服务器 （2）用户可以通过基本 ACL 和基本 ACL6、高级 ACL 和高级 ACL6 实现对 VTY 用户界面的呼入/呼出进行限制 （3）用户可以通过基本 ACL 和基本 ACL6 限定指定地址的网管终端管理设备，以及限定网管终端管理的 MIB 节点，可以增强网管终端和被管理设备使用 SNMP 进行通信时的安全性
高级 ACL 和高级 ACL6		
基本 ACL	路由策略中的应用	当需要控制路由设备接收、发布的路由信息时，可以通过配置基本 ACL 和高级 ACL 实现只接收或发布满足过滤条件的路由，实现网络路由规模的优化和网络的安全。有关路由方面的内容将在配套图书《华为路由器学习指南》一书中介绍
高级 ACL		
基本 ACL 和基本 ACL6	组播过滤中的应用	当需要过滤组播报文时，可以通过配置基本 ACL 和基本 ACL6、高级 ACL 和高级 ACL6 实现只接收或转发满足过滤条件的组播报文，从而保证了网络的可靠性和安全性
高级 ACL 和高级 ACL6		
基本 ACL 和基本 ACL6	QoS 中的应用	当需要对不同类型的流量进行分类操作时，可以通过配置基本 ACL 和基本 ACL6、高级 ACL 和高级 ACL6、二层 ACL、用户自定义 ACL 实现对满足过滤条件的流量进行 QoS 流量监管、流量整形、流分类。有关 QoS 方面将在下章具体介绍
高级 ACL 和高级 ACL6		
二层 ACL		
用户自定义 ACL		
基本 ACL	安全中的应用	当需要对某类报文进行处理，或直接丢弃时，可以通过配置基本 ACL、高级 ACL、二层 ACL 实现对攻击本机的报文进行丢弃或包过滤等功能
高级 ACL		
二层 ACL		

9.1.2 ACL编号和命名规则

前面说了，ACL的命名方式分数字型ACL和命名型ACL两种。数字型ACL就是用户在创建 ACL 时必须为其指定编号，系统将根据用户所指定的编号来创建不同的 ACL。但这里的 ACL 编号是不能随便指定的，一定要在对应类型的 ACL 的编号取值范围中，它代表了ACL的类型。如基本ACL和基本ACL6的编号取值范围都是2 000～2 999，而高级ACL和高级ACL6的编号取值范围都是 3 000～3 999，具体参见表9-1。

有时为了方便对具体ACL用途的识别，用户在创建ACL时可以为ACL指定一个名称，这就是前面所说的命名型ACL。命名型ACL可使用户通过ACL名称唯一地标识一个ACL，并对其进行相应的操作。但每个ACL最多只能有一个名称，且在**ACL**创建后不允许用户修改或者删除**ACL**名称，也不允许为未命名的**ACL**添加名称，这点要特别注意。ACL 的名称对于 ACL 全局唯一，但允许基本 ACL 与基本 ACL6，高级ACL与高级ACL6使用相同的名称。

另外要注意，在指定命名型 **ACL** 时也可以同时配置对应编号 。如果没有配置对应编号，系统在记录此命名型ACL时会自动为其分配一个数字型ACL的编号。所以命名型的 ACL也肯定有一个 ACL编号，但数字型 ACL却不会有对应的 ACL名称。

9.1.3 ACL规则编号

一个 ACL 内可以有一条或者多条规则，每条规则都有自己的编号，这是系统在进行规则匹配的缺省匹配顺序。且要求每个规则的编号在整个 ACL 中是唯一的。在创建规则时，可以人为地为每个规则指定一个唯一的编号，也可以由系统为其自动分配一个唯一的编号。当然，它们也不是随意编排的，下面具体介绍ACL规则编号的编号规则。

1．自动分配规则编号

在自动分配规则编号时，为了方便后续在已有规则之前插入新的规则（用以控制规则的匹配顺序，这

点对于想要修改**ACL**的规则匹配结果时很重要），系统通常会在相邻编号之间留下一定的空间，这个空间的大小（即相邻编号之间的差值）就称为ACL的步长。在定义一条ACL规则的时候，如果用户不指定规则编号，系统就会从现有规则中最大的ACL规则号（最小的规则号为**0**）开始，按照步长设置自动为当前添加的规则分配一个大于现有规则最大编号的最小编号。假设现有规则中的最大规则号是25，步长是5，那么系统分配给新定义的规则的编号将是30。

2．插入新规则时的规则编号

如果想要在原来的两规则之间插入一条新的规则，则插入的这条规则的编号必须手工指定，且其编号必须位于原来两条规则编号之间。假设已配置好了4个规则，规则编号为 5、10、15、20，此时如果用户希望在第一条规则之后插入一条规则，则可以使用命令在5和10之间插入一条编号为7的规则。

3．新步长的应用

ACL规则的步长既可采用系统的缺省值5，又可手工设置（但在华为S系列交换机中的基本**ACL6**和高级**ACL6**中不支持手工设定步长值）。当步长改变后，ACL中的规则编号会自动从新的步长值开始重新排列。例如，原来规则编号为5、10、15、20，当通过step step命令（本章后面具体介绍）把步长改为2后，则规则编号变成2、4、6、8。当使用 undo step命令将步长恢复为缺省值后，设备将立刻按照缺省步长调整ACL规则的编号。例如，ACL 3001，原步长设置为 2，下面有 4个规则，编号为2、4、6、8；如果此时将步长恢复为缺省值，则ACL规则编号变成5、10、15、20，步长为5。

9.1.4 ACL规则的匹配顺序

一个ACL可以由多条“deny |permit”语句组成，每一条语句描述一条规则。由于每条规则中的报文匹配选项不同（同一**ACL**中的各条规则间都不可能完全相同），从而使这些规则之间可能存在交叉甚至矛盾的地方，因此，在将一个报文与ACL的各条规则进行匹配时，就需要有明确的匹配顺序来确定规则执行的优先级。

华为 S 系列交换机的 ACL 规则匹配顺序有“配置顺序”和“自动排序”两种。当将一个数据包与访问控制列表的规则进行匹配的时候，由规则的匹配顺序设置决定规则的优先级（并不一定就是严格按照规则号大小顺序）。ACL 通过设置规则的优先级来处理规则之间重复或矛盾的情形。

（1）配置顺序：是按照用户配置规则编号的大小顺序进行匹配。我们可利用这一特点在原来规则前、后或者中间插入新的规则，以修改原来的规则匹配结果。因此，后插入的规则如果编号较小也有可能先被匹配。缺省采用配置顺序进行匹配 。

（2）自动排序：是按照“深度优先”原则由深到浅进行匹配。“深度优先”即根据规则的精确度排序，匹配条件（如协议类型、源和目的 IP 地址范围等）限制越严格越精确，优先级越高。例如可以比较地址的通配符掩码（wildcard，每位也是由“1”和“0”组成，**“0”**表示要精确匹配的位，**“1”**表示不需要匹配的位），通配符越小（“0”的位数越多），则指定的主机的范围就越小（通配符全为 0 时则表示要精确匹配地址中的每一位，相当于只有一个地址符合匹配条件，所以说主机地址的通配符为0），限制就越严格。若“深度优先”的顺序相同，则匹配该规则时按规则编号从小到大排列。

说明

通配符掩码与反向掩码类似（不完全相同），以点分十进制表示，并用二进制的“0”表示“需要进行匹配操作”，“1”表示“不需要进行匹配操作，即忽略”，这恰好与子网掩码的表示方法相反。另外通配符中的“1”或者“0”可以不连续，但掩码与反掩码必须连续。比如，IP地址192.168.1.169、通配符0.0.0.172表示的网址为192.168.1.x0x0xx01，其中x可以是0，也可以是1，但反掩码却不能有这样的值。

不同类型ACL的“深度优先”排序规则如表9-3所示。但无论是哪种匹配顺序，当报文与各条规则进行匹

配时，一旦匹配上某条规则，都不会再继续匹配下去，系统将依据该规则对该报文执行相应的操作。所以说，每个报文实际匹配的规则只有一条 。华为的ACL在最后都有一条**permit any any**，即允许所有报文通过的规则，当前面所有规则都匹配不上时将直接采用最后这条规则，允许通过。

表9-3 各类型ACL的“深度优先”排序法则

ACL 类型	“深度优先”排序规则
基于接口的 ACL	配置了 **any** 的规则排在后面，其他的规则根据规则编号大小的顺序，小的优先
基本 ACL	（1）先看规则中是否带 VPN 实例，带 VPN 实例的规则优先 （2）再比较源 IP 地址范围，源 IP 地址范围小（子网掩码中“1”位的数量多，匹配更精确）的规则优先 （3）如果源 IP 地址范围相同，则规则编号小的优先

（续表）

ACL 类型	“深度优先”排序规则
高级 IPv4 A CL	（1）先看规则中是否带 VPN 实例，带 VPN 实例的规则优先 （2）再比较所指定的协议范围，指定了 IP 协议承载的协议类型的规则优先 （3）如果协议范围相同，再比较源 IP 地址范围，源 IP 地址范围小（子网掩码中“1”位的数量多）的规则优先 （4）如果协议范围、源 IP 地址范围都相同，再比较目的 IP 地址范围，目的 IP 地址范围小（子网掩码中“1”位的数量多）的规则优先 （5）如果协议范围、源 IP 地址范围、目的 IP 地址范围都相同，再比较四层端口号（TCP/UDP 端口号）范围，四层端口号范围小的规则优先 （6）如果上述范围都相同，则规则编号小的优先
二层 ACL	（1）先比较二层协议类型通配符，通配符大（协议类型掩码（*type-mask*）中“1”位的数量多）的规则优先 （2）如果二层协议类型通配符相同，再比较源 MAC 地址范围，源 MAC 地址范围小（掩码中“1”位的数量多）的规则优先 （3）如果源 MAC 地址范围也相同，再比较目的 MAC 地址范围，目的 MAC 地址范围小（掩码中“1”位的数量多）的规则优先 （4）如果源 MAC 地址范围、目的 MAC 地址范围都相同，则规则编号小的优先
用户自定义 ACL	用户自定义 ACL 规则的匹配顺序只支持配置顺序，即按规则编号从小到大的顺序进行匹配

9.2 ACL配置

本节介绍基本ACL、高级ACL、二层ACL和用户自定义ACL的配置方法。所有ACL的配置任务很简单，主要包括以下几项配置任务。

1．配置ACL的生效时间段（可选）

时间段用于描述一个 ACL 发生作用的特殊时间范围。用户可能有这样的需求，即一些ACL规则需要在某个或某些特定时间内生效，而在其他时间段不生效。例如某单位严禁员工上班时间浏览非工作网站，而下班后则允许通过指定设备浏览娱乐网站，就可以对ACL规则约定生效时间段。这时用户就可以先配置一个或多个时间段，然后通过配置规则引用该时间段，从而实现基于时间段的ACL过滤。但如果规则中引用的时间段未配置，则整个规则不能立即生效 ，直到用户配置了引用的时间段，并且系统时间在指定时间段范围内，ACL规则才能生效。

时间段的配置包括以下两种方式。

（1）相对时间段（周期时间段）：采用每个星期固定时间段的形式，例如从星期一到星期五的8:00至18:00。

（2）绝对时间段：采用从某年某月某日某时某分起至某年某月某日某时某分结束的形式，例如从2011年4月28日10:00起至2012年4月28日10:00结束。

2．创建ACL

配置ACL时需要先创建一个基本ACL，可以是数字型的，也可以是命名型的。如果是数据型的，其编号一定要在对应类型的 ACL编号范围之内，如基本 ACL编号的取值范围为 2 000～2 999，高级ACL的编号取值范围为 3 000～3 999，二层ACL的编 号取值范围为 4 000～4 999，用户自定义ACL的编号取值范围为 5 000～5 999，一定不能用错。

3．配置ACL规则

ACL 通过具体的规则（rule）所指定的过滤条件来匹配报文中的信息，实现对报文的分类。因此，创建ACL以后，需要根据不同类型ACL可以匹配的参数配置满足对应需求的各条ACL规则。

在 ACL 中添加新的规则时，不会影响已经存在的规则（但可能会改变原有规则的匹配顺序 ）；对已经存在的规则进行编辑时，如果新配置的规则内容与原规则内容存在冲突，则冲突的部分由新配置的规则内容代替（通过这种方式也可修改原来的规则 ）。但建议在编辑一个已存在的规则前，先删除旧的规则，再创建新的规则，否则配置结果可能与预期的效果不同。此外，配置规则时如果不同的规则之间存在矛盾或包含的关系，则要充分考虑前面在9.1.4节所介绍的规则匹配顺序，以防错误配置。

说明

ACL配置好后还需要在对应位置或者功能中应用，在华为S系列交换机中，多数 情况下是通过基于ACL的简化流策略的方式在交换机所有端口/所有VLAN，或者指定VLAN、指定端口的出或入方向上应用，具体将在本章后面介绍。至于其他像路由信息过滤、策略路由等方面的应用将在《华为路由器学习指南》一书中介绍。

9.2.1 配置基本ACL

基本ACL是应用于IPv4协议网络的，且基本ACL的地址过滤信息中一定不会包括目的 IP 地址，只包括源 IP 地址，这是基本 ACL 的最显著特点。根据前面介绍的配置ACL的三项任务，可得出的具体配置步骤如表9-4所示。

表9-4 基本ACL的配置步骤

配置任务	步骤	命令	说明
公共配置步骤	1	**system-view** 例如：<HUAWEI> **system-view**	进入系统视图
（可选）配置ACL 生效时间段	2	**time-range** *time-name* { *start-time* **to** *end-time* *days* | **from** *time1 date1* [**to** *time2 date2*] } 例如：[HUAWEI] **time-range** test 14:00 **to** 18:00 **off-day**	创建一个指定 ACL 生效的时间段。命令中的参数说明如下。 （1）*time-name*：定义时间段的名称，作为一个引用时间段的标识。为 1～32 个字符的字符串，**区分大小写**，但必须以英文字母 a～z 或 A～Z 开头。为避免混淆，时间段的名称不允许使用英文单词 all，且同一名称时间段下面可以配置多个不同的时间段 （2）*start-time* **to** *end-time*：二选一参数，指定周期时间段的时间范围，参数 *start-time* 和 *end-time* 分别表示起始时间和结束时间，格式均为 hh:mm（小时：分钟）。hh 的取值范围为 0～23，mm 的取值范围为 0～59，且结束时间必须大于起始时间 （3）*days*：与上面的“*start-time* **to** *end-time*”参数

（续表）

配置任务	步骤	命令	说明	
（可选） 配置 ACL 生效 时间段	2	**time-range** *time-name* { *start-time* **to** *end-time* *days* | **from** *time1 date1* [**to** *time2 date2*] } 例如：[HUAWEI] **time-range** test 14:00 **to** 18:00 **off-day**	一起构成一个二选一参数，指定周期时间段在每周的周几生效。有如下输入格式。 ① 0～6 数字的表示周日期，其中 0 表示星期天。此格式支持输入多个参数，各个值之间以空格分配 ② **Mon**、**Tue**、**Wed**、**Thu**、**Fri**、**Sat**、**Sun** 英文表示的周日期，分别对应星期一到星期日。此格式支持输入多个参数，各个值之间以空格分配 ③ **daily** 表示所有日子，包括一周共 7 天 ④ **off-day** 表示休息日，包括星期六和星期天 ⑤ **working-day** 表示工作日，包括从星期一到星期五 （4）*time1 date1*：二选一参数，指定绝对时间段的开始日期，表示从某一天某一时间开始。它的表示形式为 hh:mm YYYY/MM/DD（小时：分钟 年/月/日）或 hh:mm MM/DD/YYYY（小时：分钟 月/日/年） （5）*time2 date2*：可选参数，指定绝对时间段的结束日期，表示到某一天某一时间结束。它的表示形式也为 hh:mm YYYY/MM/DD 或 hh:mm MM/DD/ YYYY 缺省情况下，设备没有配置时间段，可用 **undo time-range** *time-name* [*start-time* **to** *end-time* { *days* } &<1-7> | **from** *time1 date1* [**to** *time2 date2*]]命令删除一个指定的时间段，或者指定名称下的所有时间段（当不选择所有可选参数时）	
创建基本 ACL	3	**acl** [**number**] *acl-number* [**match-order** { **auto** | **config** }] 例如：[HUAWEI] **acl number** 2100	使用编号创建一个数字型的基本 ACL，并进入基本 ACL 视图。命令中的参数和选项说明如下。 （1）**number**：可选项，指定创建数字型 ACL，缺省也是数字型的，所以也可以不选择此可选项 （2）*acl-number*：用来指定基本 ACL 的编号，取值范围为 2 000～2 999 （3）**match-order** { **auto** | **config** }：可选项，用来指定规则的匹配顺序。二选一选项 **auto** 表示按照自动排序（即按“深度优先”原则）的顺序进行规则匹配，若“深度优先”的顺序相同，则匹配规则时按规则号由小到大的顺序；二选一选项 **config** 表示按照配置顺序进行规则匹配，即在用户没有指定规则编号时按用户的配置顺序进行匹配；如果用户指定了规则编号，则按规则编号由小到大的顺序进行匹配。缺省情况下，规则的匹配顺序为配置顺序。**但仅 S7700/9300/9700 系列交换机支持** 缺省情况下，不存在任何 ACL，可用 **undo acl** { [**number**] *acl-number* | **all** }命令删除指定的，或者所有基本 ACL。删除 ACL 时，如果删除的 ACL 被其他业务引用，可能造成该业务的中断，所以在删除 ACL 时请先确认是否有业务正在引用该 ACL	二选一

（续表）

配置任务	步骤	命令	说明	
创建基本 ACL	3	**acl name** *acl-name* { **basic** | *acl-number* } [**match-order** { **auto** | **config** }] 例如：[HUAWEI] **acl name** test1 2001	使用名称创建一个命名型的基本 ACL，并进入基本 ACL 视图。命令中的参数和选项说明如下： （1）*acl-name*：指定创建的 ACL 的名称，为 1～32 个字符，**区分大小写**，且需以英文字母 a～z 或 A～Z 开始 （2）**basic**：二选一选项，指定 ACL 的类型为基本 ACL。此时设备为其分配的 ACL 编号是该类型 ACL 可用编号中取值范围内的最大值。设备不会为命名型 ACL 重复分配编号 （3）*acl-number*：二选一选项，指定基本 ACL 的编号，取值范围为 2 000～2 999 （4）**match-order** { **auto** | **config** }：可选项，用来指定规则的匹配顺序。具体说明同上一步 缺省情况下，系统中没有创建命名型 ACL，可用 **undo acl name** *acl-name* 来删除指定的命名型 ACL	二选一
	4	**description** *text* 例如：[HUAWEI-acl-basic-2100] **description** This acl is used in Qos policy	（可选）定义 ACL 的描述信息，主要目的是便于理解，比如可以用来描述该 ACL 规则列表的具体用途。参数 *text* 表示 ACL 的描述信息，为 1～127 个字符的字符串，**也区分大小写** 缺省情况下，ACL 没有描述信息，可用 **undo description** 命令删除 ACL 的描述信息	
	5	**step** *step* 例如：[HUAWEI-acl-basic-2100] **step** 8	（可选）为一个 ACL 规则组中的规则编号配置步长，取值范围是 1～20 的整数。缺省情况下，步长值为 5，可用 **undo step** 命令用来恢复规则编号的步长为缺省值	
配置基本 ACL 规则	6	**rule** [*rule-id*] { **deny** | **permit** } [**source** { *source-address source-wildcard* | **any** } | **fragment** | **logging** | **time-range** *time-name*] * 例如：[HUAWEI-acl-basic-2100] **rule permit source** 192.168.32.1 0	配置基本 ACL 的规则。**如果需要配置多个规则，可以反复执行本命令**。命令中的参数和选项说明如下。 （1）*rule-id*：可选参数，用来指定基本 ACL 规则的编号，取值范围为 0～4 294 967 294 的整数。如果指定规则号的规则已经存在，则会在旧规则的基础上叠加新定义的规则，相当于编辑一个已经存在的规则；如果指定的规则号的规则不存在，则使用指定的规则号创建一个新规则，并且按照规则号的大小决定规则插入的位置。 如果不指定本参数，则增加一个新规则时设备自动会为这个规则分配一个规则号，规则号按照大小排序。系统自动分配规则号时会留有一定的空间，相邻规则号的范围由上一步的 **step** *step* 命令指定 （2）**deny**：二选一选项，设置拒绝型操作，表示拒绝符合条件的报文通过 （3）**permit**：二选一选项，设置允许型操作，表示允许符合条件的报文通过 （4）**source** { *sour-addr sour-wildcard* | **any** }：可多选项，指定规则的源地址信息。二选一参数 *sour-addr sour-wildcard* 分别表示报文的源 IP 地址和通配符（通配符是用来确定对应位是否要匹配的，值为“0”的位表示要匹配，值为“1”的位表示不需要匹配，当全为 0 时表示源 IP 地址为主机地址，表示 IP 地址中	

（续表）

配置任务	步骤	命令	说明
配置基本 ACL 规则	6	**rule** [*rule-id*] { **deny** | **permit** } [**source** { *source-address source-wildcard* | **any** } | **fragment** | **logging** | **time-range** *time-name*] * 例如：[HUAWEI-acl-basic-2100] **rule permit source** 192.168.32.1 0	的每一位都需要匹配）；二选一选项 **any** 表示任意源 IP 地址，相当于 *source-address* 为 0.0.0.0 或者 *source-wildcard* 为 255.255.255.255 （5）**fragment**：可多选项，表示该规则仅对非首片分片报文有效，而对非分片报文和首片分片报文无效。如果没有指定本参数，则表示该规则对非分片报文和分片报文均有效 （6）**logging**：可多选项，指定将该规则匹配的报文的 IP 信息进行日志记录 （7）*time-range-name*：可多选项，指定该规则生效的时间段。*time-range-name* 表示时间段的名称，为 1～32 个字符的字符串，**区分大小写**，但必须以英文字母 a～z 或 A～Z 开头 缺省情况下，未配置任何规则，可用 **undo rule** *rule-id* [**fragment** | **logging** | **source** | **time-range**] * 命令在对应 ACL 视图下删除指定的一条规则或一条规则中的部分内容
	7	**rule** *rule-id* **description** *description* 例如：[HUAWEI-acl-basic-2001] **rule 5 description** permit 192.168.32.1	（可选）配置基本 ACL 规则的描述信息。命令中的参数说明如下。 （1）*rule-id*：指定要描述的 ACL 规则的编号，取值范围为 0～4 294 967 294 的整数 （2）*description*：指定某规则号的规则描述信息。用户可以通过这个描述信息更详细地记录规则，便于识别规则的用途，为 1～127 个字符

使用在配置 ACL 生效时间段时，配置的时间段时间不能早于当前时间，否则不会生效。另外，可以为

多个时间段配置相同的时间段名称，共同来描述某个特殊时间。通过名称来引用一个时间段时，如果该时间段配置了多个生效时间，生效原则为各周期时间段之间以及各绝对时间段之间分别取并集之后，再取二者的交集作为最终生效的时间范围。例如：时间段“test”配置了三个生效时段。

（1）从2013年6月1日00:00起到2014年5月31日23:59生效，这是一个绝对时间段。

（2）在周一到周五每天8:00到18:00生效，这是一个周期时间段。

（3）在周六、周日下午14:00到18:00生效，这是一个周期时间段。

则“test”最终将在以下时间内生效：2013年6月1日起到2014年5月31日23:59内的周一到周五每天8:00到18:00以及周六和周日下午14:00到18:00。

【示例 1】配置时间段test，从2013年1月1日00:00起到2013年12月31日23:59生效。

<HUAWEI>system-view

[HUAWEI] time-range test from 0:0 2013/1/1 to 23:59 2013/12/31

【示例 2】配置时间段test，在周一到周五每天8:00到18:00生效。

<HUAWEI>system-view

[HUAWEI] time-range test 8:00 to 18:00working-day

【示例 3】配置时间段test，在周六、周日下午14:00到18:00生效。

<HUAWEI>system-view

[HUAWEI] time-range test 14:00 to 18:00 off-day

【示例 4】配置在ACL 2001中增加一条规则，允许源地址是 192.168.32.1的报文通过。

<HUAWEI>system-view

[HUAWEI] acl 2001

[HUAWEI-acl-basic-2001] rule permit source 192.168.32.1 0

【示例 5】配置在ACL 2001中删除一条规则，删除规则 5。

<HUAWEI>system-view

[HUAWEI] acl 2001

[HUAWEI-acl-basic-2001] undo rule 5

9.2.2 配置高级ACL

高级ACL除了可以根据上节介绍的基本ACL中的报文源IP 地址进行规则匹配之外，还可以根据报文的目的IP地址信息、IP承载的协议类型、协议的特性（如TCP或UDP的源端口、目的端口，ICMP协议的消息类型、消息码等）等信息进行匹配。

另外，高级ACL还支持QoS中所需的优先级设置，用于指定符合对应优先级的IP数据包通过。当用户需要使用源IP地址、目的IP地址、源端口号、目的端口号、优先级、时间段等信息对 IPv4 报文进行过滤时，可以使用高级 ACL。高级 ACL 主要用于QoS中的流分类，因为通过它可以精确地对流量进行分类。

说明

在IPv4网络中，IPv4报文中有三种承载QoS优先级标签的方式：基于二层的CoS （Class of Service，服务等级）字段（即通常所说的 802.1p优先级、基于 IP层（三层）的IP优先级字段（即IP优先级）和ToS（服务类型）字段，以及基于IP层（三层）的DSCP（Differentiated Services Code Point，差分服务代码点）字段（即DSCP优先级））。华为S系列交换机中的高级ACL支持ToS优先级、IP优先级和DSCP优先级这三种优先级。具体参见本书第10章10.1.2和10.1.3节介绍。

在前面介绍的ACL配置任务中，ACL生效时间段的配置方法与上节基本ACL配置中的时间段配置方法

完全一样，下面仅介绍不同的高级ACL创建和ACL规则配置两个方面。

1．高级ACL的创建

在高级ACL的创建中，同样可以创建数字型的，或者命名型的ACL。如果创建的是数字型高级ACL，则创建方法与前面介绍的基本ACL的创建方法一样，使用的也是acl [number] acl-number [match-order { auto | config }] 命令，但其中 acl-number参数的取值范围只能是 3 000～3 999，同样[match-order { auto | config }]可选项仅 S7700/9300/9700系列交换机支持。

如果创建的是命名型高级 ACL，则需要采用 acl name acl-name {advance |acl-number } [match-order { auto | config }] 命令，选择二选一选项 advance时表示所创建的是命名型高级 ACL，选择二选一参数 acl-number时，则其取值范围也是 3 000～3 999，表示创建的是命名型高级ACL。

2．高级ACL规则的配置

高级ACL规则的配置比基本ACL中的规则配置复杂许多，因为可用来匹配的过滤 条件参数非常之多，而且基本上是可同时配置的。根据IP包中承载的协议类型不同，在设备上可配置提供不同匹配条件参数的高级 ACL，具体如下（注意其中的"*"表示前面用"[ ] "括住的参数或选项是可同时多选的）：

（1）当参数protocol为ICMP时（即要过滤ICMP协议报文时），命令格式为：

rule [rule-id] {deny |permit } {protocol-number | icmp } [destination {destination- address destination-wildcard | any } | { {precedence precedence | tos tos } * |dscpdscp } | fragment | logging | icmp-type { icmp-name | icmp-type icmp-code } |source { source-address source-wildcard | any } | time-range time-name | ttl-expired] *

（2）当参数protocol为TCP时（即要过滤TCP协议报文时），命令格式为：

rule [rule-id] {deny |permit } { protocol-number | tcp } [destination {destination- address destination-wildcard | any } |destination-port { eq port | gt port | ltport | range port-startport-end } | { {precedenceprecedence | tos tos} * |dscp dscp } | fragment | logging |source { source-address source-wildcard |any } | source-port {eqport |gtport | lt port | range port-startport-end } | tcp-flag {ack | fin |psh | rst | syn |urg }* | time-range time-name | ttl-expired] *

（3）当参数protocol为UDP时（即要过滤UDP协议报文时），命令格式为：

rule [rule-id] {deny |permit } {protocol-number |udp } [destination {destination- address destination-wildcard | any } |destination-port { eqport | gtport | ltport | rangeport- startport-end } | { {precedence precedence | tos tos } * |dscp dscp } | fragment | logging |source { source-address source-wildcard |any } | source-port { eqport |gtport | ltport |range port-start port-end } | time-range time-name | ttl-expired] *

（4）当参数protocol为GRE、IGMP、IP、IPINIP、OSPF时，命令格式为：

rule [rule-id] {deny |permit } {protocol-number |gre | igmp | ip | ipinip |ospf } [destination {destination-address destination-wildcard |any } | { {precedence precedence | tos tos} * |dscp dscp } | fragment | logging | source { source-address source-wildcard | any } | time-range time-name | ttl-expired] *

缺省情况下，未配置ACL规则，可用undo rule rule-id [destination |destination-port | dscp | fragment | logging | icmp-type | precedence | source | source-port | tcp-flag |time-range | tos | ttl-expired] *删除一个指定的ACL规则。

以上rule命令中的参数和选项说明如表9-5所示。

表9-5 rule命令参数和选项说明

参数	说明
rule-id	可选参数，用来指定高级 ACL 规则的编号，取值范围为 0～4 294 967 294 的整数。其他说明参见表 9-4 中的第 6 步
deny	二选一选项，表示拒绝符合条件的报文
permit	二选一选项，表示允许符合条件的报文
icmp	二选一选项，指定 ACL 规则匹配报文的协议类型为 ICMP。也可以通过参数 *protocol-number* 采用数值 1 表示指定 ICMP 协议
tcp	二选一选项，指定 ACL 规则匹配报文的协议类型为 TCP。也可以通过参数 *protocol-number* 采用数值 6 表示指定 TCP 协议

（续表）

参数	说明
udp	二选一选项，指定 ACL 规则匹配报文的协议类型为 UDP。也可以通过参数 *protocol-number* 采用数值 17 表示 UDP 协议
gre	多选一选项，指定 ACL 规则匹配报文的协议类型为 GRE。也可以通过参数 *protocol-number* 采用数值 47 表示 UDP 协议
igmp	多选一选项，指定 ACL 规则匹配报文的协议类型为 IGMP。也可以通过参数 *protocol-number* 采用数值 2 表示 UDP 协议
ip	多选一选项，指定 ACL 规则匹配报文的协议类型为 IP
ipinip	多选一选项，指定 ACL 规则匹配报文的协议类型为 **ipinip**。也可以通过参数 *protocol-number* 采用数值 4 表示 UDP 协议
ospf	多选一选项，指定 ACL 规则匹配报文的协议类型为 **OSPF**。也可以通过参数 *protocol-number* 采用数值 89 表示 UDP 协议
protocol-number	多选一参数，用数字表示报文中封装的协议类型，取值范围为 1～255，如 **tcp** 是 6，**udp** 是 17，**icmp** 是 1，**tcp** 是 6，**udp** 是 17，**gre** 是 47，**igmp** 是 2，**ipinip** 是 4，**ospf** 是 89
source { *sour-addr sour-wildcard* | **any** }	可多选参数，指定 ACL 规则的源 IP 地址信息。二选一参数对 *sour-addr sour-wildcard* 是用来指定源 IP 地址及其通配符掩码（通配符为 0 时表示源地址为主机 IP 地址）；二选一选项 **any** 表示为任意源 IP 地址
destination { *dest-addr dest-wildcard* | **any** }	可多选参数，指定 ACL 规则的目的 IP 地址信息。二选一参数对 *dest-addr dest-wildcard* 用来指定目的 IP 地址及其通配符掩码（通配符为 0 时表示目的地址为主机 IP 地址）；二选一选项 **any** 表示任意目的 IP 地址
icmp-type { *icmp-name* | *icmp-type icmp-code* }	可多选参数，指定 ACL 规则匹配报文的 ICMP 报文的类型和消息码信息，仅在报文协议是 ICMP 的情况下有效。如果不配置，表示任何 ICMP 类型的报文都匹配。其中： （1）*icmp-name*：表示 ICMP 的消息名称 （2）*icmp-type*：表示 ICMP 的消息类型，取值范围为 0～255 的整数 （3）*icmp-code*：表示 ICMP 的消息码，取值范围为 0～255 的整数 参数 *icmp-name* 的取值与参数 *icmp-type* 和 *icmp-code* 相对应，对应关系如表 9-6 所示
precedence *precedence*	可多选参数，指定 ACL 规则匹配报文时依据优先级字段进行过滤。它与 **tos** *tos* 参数一起共同构成与 **dscp** *dscp* 组成的二选一参数 参数 *precedence* 表示 IP 优先级字段值。参数 *precedence* 用数字表示时，取值范围为 0～7；用名称表示时对应为 **routine**、**priority**、**immediate**、**flash**、**flash-override**、**critical**、**internet control**、**network control**
tos *tos*	可多选参数，指定 ACL 规则匹配报文时，依据服务类型字段进行过滤。它与 **precedence** *precedence* 参数一起共同构成与 **dscp** *dscp* 组成的二选一参数 参数 *tos* 表示 IP 服务等级，用数字表示时取值范围为 0～15（但只能是 0、1、2、4、8 这五个数）；用名称表示时对应为 **normal**、**min-monetary-cost**、**max-reliability**、**max-throughput**、**min-delay**
dscp *dscp*	二选一参数，指定 ACL 规则匹配报文时，区分服务代码点，即依据 IP 包中的 DSCP 优先级字段进行过滤。参数 *dscp* 用数字表示时，取值范围为 0～63（但不是全部可选，参照下面以名称表示时对应的值）；用名称表示时，可选取 **af11**（10）、**af12**（12）、**af13**（14）、**af21**（18）、**af22**（20）、**af23**（22）、**af31**（26）、**af32**（28）、**af33**（30）、**af41**（34）、**af42**（36）、**af43**（38）、**cs1**（8）、**cs2**（16）、**cs3**（24）、**cs4**（32）、**cs5**（40）、**cs6**（48）、**cs7**（56）、**default**（0）或 **ef**（46）

（续表）

参数	说明
tcp-flag { **ack** | **fin** | **psh** | **rst** | **syn** | **urg** }*	可多选项，指定 ACL 规则匹配 TCP 报文头中 SYN 标志字段的类型依次为 ack（010000）、fin（000001）、psh（001000）、rst（000100）、syn（000010））和 urg（100000）。注意，括号中的数字代表选择对应标志位时的 SYN 标志字段值，不用输入
time-range *time-range-name*	可多选参数，指定 ACL 规则生效的时间段。参数 *time-range-name* 用来设置时间段的名称，为 1～32 个字符的字符串，区分大小写，必须以英文字母 a～z 或 A～Z 开头。如果不指定时间段，表示任何时间都生效
destination-port { **eq** *port* | **gt** *port* | **lt** *port* | **range** *port-start port-end* }	可多选参数，指定 ACL 规则匹配报文的 UDP 或者 TCP 目的端口，仅在报文协议是 TCP 或者 UDP 时有效。如果不指定，表示 TCP/UDP 报文的任何目的端口都匹配。其中： （1）**eq** *port*：指定等于目的端口 （2）**gt** *port*：指定大于目的端口 （3）**lt** *port*：指定小于目的端口 （4）**range** *port-start port-end*：指定目的端口的范围：*port-start* 是目的端口范围的起始端口号，*port-end* 是目的端口范围的结束端口号 以上的参数 *port* 用来指定端口号，用名字或数字表示。用数字表示时，**eq** *port*、**lt** *port*、*port-start* 和 *port-end* 中的 *port* 的取值范围是 0～65 535；**gt** *port* 中的 *port* 的取值范围是 0～65 534
	用名称表示时，TCP 端口号可选取 **chargen**（19）、**bgp**（179）、**cmd**（514）、**daytime**（13）、**discard**（9）、**domain**（53）、**echo**（7）、**exec**（512）、**finger**（79）、**ftp**（21）、**ftp-data**（20）、**gopher**（70）、**hostname**（101）、**irc**（194）、**klogin**（543）、**kshell**（544）、**login**（513）、**lpd**（515）、**nntp**（119）、**pop2**（109）、**pop3**（110）、**smtp**（25）、**sunrpc**（111）、**tacacs**（49）、**talk**（517）、**telnet**（23）、**time**（37）、**uucp**（540）、**whois**（43）或 **www**（80）；UDP 端口号可选取 **biff**（512）、**bootpc**（68）、**bootps**（67）、**discard**（9）、**dns**（53）、**dnsix**（90）、**echo**（7）、**mobilip-ag**（434）、**mobilip-mn**（435）、**nameserver**（42）、**netbios-dgm**（138）、**netbios-ns**（137）、**netbios-ssn**（139）、**ntp**（123）、**rip**（520）、**snmp**（161）、**snmptrap**（162）、**sunrpc**（111）、**syslog**（514）、**tacacs-ds**（65）、**talk**（517）、**tftp**（69）、**time**（37）、**who**（513）或 **xdmcp**（177）。注意，括号中的数字对应的端口号，不用输入
source-port { **eq** *port* | **gt** *port* | **lt** *port* | **range** *port-start port-end* }	可多选参数，指定 ACL 规则匹配报文的 UDP 或者 TCP 报文的源端口，仅在报文协议是 TCP 或者 UDP 时有效。如果不指定，表示 TCP/UDP 报文的任何源端口都匹配。其他说明参见 **destination-port** 选项中的对应说明
logging	可多选项，指定将该规则匹配的报文的 IP 信息记录日志
fragment	可多选项，指定该规则是否仅对非首片分片报文有效。当包含此参数时表示该规则仅对非首片分片报文有效。但 **fragment** 选项不能跟 **source-port**、**destination-port**、**icmp-type** 和 **tcp-flag** 参数同时配置，否则会提示错误信息
ttl-expired	可多选项，数据包可以依据报文中的 TTL 字段值是否为 1 进行过滤，如果不配置，表示在过滤报文时不考虑 TTL 字段值。但 **S2700/3700/5700SI/5700LI/5700S-LI 系列交换机不支持本**可选项

表9-6 ICMP消息名称与消息类型和消息码的对应关系

icmp-name（ICMP 消息名称）	icmp-type（ICMP 消息类型）	icmp-code（ICMP 消息代码）
Echo	8	0
Echo-reply	0	0
Parameter-problem	12	0

（续表）

icmp-name（ICMP 消息名称）	icmp-type（ICMP 消息类型）	icmp-code（ICMP 消息代码）
Port-unreachable	3	3
Protocol-unreachable	3	2
Reassembly-timeout	11	1
Source-quench	4	0
Source-route-failed	3	5
Timestamp-reply	14	0
Timestamp-request	13	0
Ttl-exceeded	11	0
Fragmentneed-DFset	3	4
Host-redirect	5	1
Host-tos-redirect	5	3
Host-unreachable	3	1
Information-reply	16	0
Information-request	15	0
Net-redirect	5	0
Net-tos-redirect	5	2
Net-unreachable	3	0

【示例 1】配置在ACL3000（采用缺省步长5）中增加一条规则号为2的，允许基于IGMP协议类型报文通过的规则。

<HUAWEI>system-view

[HUAWEI] acl 3000

[HUAWEI-acl-adv-3000] rule 2permit igmp

【示例 2】配置在ACL3000中删除上一示例中基于ICMP协议类型的规则。

<HUAWEI>system-view

[HUAWEI] acl 3000

[HUAWEI-acl-adv-3000] undo rule 2

【示例 3】配置在ACL3001中采用规则号自动分配方式（此时不用手工指定规则号 ）增加一条规则，允许从129.9.0.0网段的主机向202.38.160.0网段的主机发送的所有IP报文通过。

<HUAWEI>system-view

[HUAWEI] acl 3001

[HUAWEI-acl-adv-3001] rule permit ip source 129.9.0.0 0.0.255.255destination 202.38.160.0 0.0.0.255

【示例 4】配置在ACL3001中采用规则号自动分配方式（此时不用手工指定规则号 ）增加一条规则，允许 129.9.8.0网段内的主机建立与 202.38.160.0网段内的主机UDP 128端口上建立的UDP通信。

<HUAWEI>system-view

[HUAWEI] acl 3001

[HUAWEI-acl-adv-3001] rule permit udp source 129.9.8.0 0.0.0.255 destination 202.38.160.0 0.0.0.255 destination-porteq 128

9.2.3 配置二层ACL

二层ACL是根据报文的源MAC地址、目的MAC地址、802.1p优先级、二层协议 类型等二层信息进行规则匹配、处理的。二层ACL的序号取值范围为 4 000～4 999。有关802.1p优先级请参见本书第10章的10.1.2节介绍。

在二层 ACL 的配置任务中第一项的时间段配置方法与 9.2.1 节中介绍的基本 ACL时间段配置方法完全一样，所以下面仅介绍二层ACL的创建和规则的配置。

1．二层ACL的创建

在二层ACL的创建中，同样可以创建数字型的，或者命名型的ACL。如果创建的是数字型二层ACL，则创建方法与前面介绍的基本ACL的创建方法一样，使用的也是acl [number] acl-number [match-order { auto | config }] 命令，但其中 acl-number参数的取值范围只能是 4 000～4 999，同样，[match-order {auto | config }] 可选项仅S7700/9300/9700系列交换机支持。

如果创建的是命名型二层ACL，则需要采用 acl nameacl-name { link | acl-number } [match-order { auto | config }] 命令，二选一选项 link用来表示所创建的是命名型二层ACL，如果选择acl-number二选一参数，则其取值范围也是4000～4999，表示创建的是命名型二层ACL。

2．二层ACL规则的配置

二层ACL规则的配置也比基本ACL中的规则配置要复杂许多，可用来匹配的过滤条件参数比较多，而且也基本上是可同时配置的。具体命令如下：

rule [rule-id] { permit | deny } [[ether-ii | 802.3 | snap] | l2-protocol type-value [type-mask] | destination-mac dest-mac-address [dest-mac-mask] | source-mac source- mac-address [source-mac-mask] | vlan-id vlan-id [vlan-id-mask] | 8021p 802.1p-value |cvlan-id cvlan-id [cvlan-id-mask] | cvlan-8021p 802.1p-value | double-tag] * [time-range time-name]

缺省情况下，未配置规则，可用undo rule rule-id命令删除一个二层ACL规则。

说明

在 S2700/3700/5700SI/5700LI/S5700S-LI 系列交换机中不支持 cvlan-id cvlan-id [cvlan-id-mask] 、cvlan-8021p802.1p-value、double-tag这几个参数或选项。

另外，参数中的地址掩码和VLN ID掩码值是十六进制，但不需要输入前面的 0x（但V2R1版本必须输入0x），直接输入后面的值即可。

rule命令中的参数和选项说明如表9-7所示。

表9-7 二层ACL规则配置命令中的参数和选项说明

参数	说明
rule-id	可选参数，用来指定二层 ACL 规则的编号，取值范围为 0～4 294 967 294 的整数。其他说明参见 9.2.1 节表 9-4 中的第 6 步
deny	二选一选项，表示拒绝符合条件的报文
permit	二选一选项，表示允许符合条件的报文
ether-ii | 802.3 | snap	可多选项，指定 ACL 规则匹配报文的帧封装格式。其中 3 个多选一选项的含义具体如下。 （1）**ether-ii** 表示匹配 Ethernet II 标准封装 （2）**802.3** 表示匹配 802.3 标准封装 （3）**snap** 表示匹配 SNAP 标准封装

（续表）

参数	说明
l2-protocol *type-value* [*type-mask*]	指定 ACL 规则匹配报文的链路层协议类型，其中：*type-value* 表示以 16 位的十六进制数标识的二层协议类型，对应 Ethernet_II 类型和 Ethernet_SNAP 类型帧中的 Type（类型）字段（2 个字节）的值，取值范围为 0x0000～0xFFFF；可选参数 *type-mask* 表示二层协议类型掩码，为 16 比特的十六进制数，用于指定屏蔽位（0 表示不需要匹配，f 表示需要匹配，**注意这里与前面说到的“通配符掩码”是相反的**），可以用来指定一个协议类型值范围，缺省值为 0xffff，即对每位都进行匹配，即仅指定一个协议类型值。常见的一些协议类型所对应的十六进制类型值如下。 0000～05DC：IEEE802.3 0800：IPv4 0805：X.25 Level 3 0806：ARP 0A00：Xerox IEEE802.3 PUP 0A01：Xerox IEEE802.3 PUP Address Translation 8035：Reverse Address Resolution Protocol (RARP) 8037：IPX 803C：DEC DNS Naming Service 803D：DEC Ethernet CSMA/CD Encryption Protocol 803E：DEC Distributed Time Service 809B：EtherTalk (AppleTalk over Ethernet) 80D5：IBM SNA Services over Ethernet 80F3：AppleTalk Address Resolution Protocol (AARP) 86DD：IPv6 9000：Loopback (Configuration Test Protocol) 9001：3Com XNS Systems Management 9002：3Com TCP/IP Systems Management 9003：3Com loopback detection AAAA：DECNET
destination-mac *dest-mac-address* [*dest-mac-mask*]	可多选项，指定 ACL 规则匹配报文的目的 MAC 地址信息。参数 *dest-mac-address* 和 *dest-mac-mask* 的格式均为 H-H-H，其中 H 为 1 至 4 位的十六进制数。其中：*dest-mac-address* 用来指定要匹配的数据包的目的 MAC 地址；可选参数 *dest-mac-mask* 用来指定目的 MAC 地址掩码，用于指定屏蔽位（0 表示不需要匹配，f 表示需要匹配），可以用来指定一个 MAC 地址范围。参数 *dest-mac-mask* 的缺省值为 ffff-ffff-ffff，即对每位都进行匹配，即仅指定一个 MAC 地址，如果不配置此可选参数，则掩码相当于 ffff-ffff-ffff 这两个参数共同作用可以定义用户想要匹配的目的 MAC 地址范围。比如 00e0-fc01-0101 ffff-ffff-ffff 指定了一个 MAC 地址：00e0-fc01-0101，而 00e0-fc01-0101 ffff-ffff-0000 则指定了一个 MAC 地址范围：00e0-fc01-0000～00e0-fc01-ffff
source-mac *source-mac-address* [*source-mac-mask*]	可多选项，指定 ACL 规则匹配报文的源 MAC 地址信息。其中：*source-mac-address*：用来指定要匹配的数据包的源 MAC 地址；可参数 *source-mac-mask* 用来指定源 MAC 地址掩码，如果不配置此可选参数，则掩码相当于 ffff-ffff-ffff 其他说明与上面介绍的 **destination-mac** *dest-mac-address* [*dest-mac-mask*]参数的说明一样

（续表）

参数	说明
vlan-id *vlan-id* [*vlan-id-mask*]	可多选项，指定 ACL 规则匹配报文的外层 VLAN 的编号，其中：*vlan-id* 用来指定要匹配的外层 VLAN ID 的值，取值范围为 1～4 094；*vlan-id-mask* 用来指定外层 VLAN ID 值的掩码（与前面介绍的 MAC 地址掩码作用一样），为十六进制形式，取值范围为 0x0～0xfff，缺省值为 0xfff（表示每位均需要匹配），可以用来指定一个外层 VLAN 或一个范围的外层 VLAN。如果不配置此参数，则掩码相当于 0xfff，即仅指定一个外层 VLAN
8021p *802.1p-value*	可多选项，指定 ACL 规则匹配报文的外层 VLAN 的 802.1p 优先级，取值范围为 0～7 的整数，对应的优先级名称为：**best-effort**、**background**、**spare**、**excellent-effort**、**controlled-load**、**video**、**voice** 和 **network-management**
cvlan-id *cvlan-id* [*cvlan-id-mask*]	可多选项，指定 ACL 规则匹配报文的内层 VLAN 的编号。其中：*vlan-id* 用来指定外层 VLAN ID 的值，取值范围是 1～4 094；*vlan-id-mask* 用来指定外层 VLAN ID 值的掩码，为十六进制形式，取值范围为 0x0～0xfff，缺省值为 0xfff，可以用来指定一个内层 VLAN 或一个范围的内层 VLAN。如果不配置此参数，则掩码相当于 0xfff，即仅指定一个内层 VLAN
cvlan-8021p *802.1p-value*	可多选项，指定 ACL 规则匹配报文的内层 VLAN 的 802.1p 优先级，取值范围为 0～7 的整数，对应的优先级名称为：**best-effort**、**background**、**spare**、**excellent-effort**、**controlled-load**、**video**、**voice** 和 **network-management**
double-tag	可多选项，指定 ACL 规则匹配报文时匹配带双层 tag 的报文
time-range *time-name*	可选参数，指定 ACL 规则生效的时间段。*time-name* 表示时间段的名称，为 1～32 个字符

【示例 1】创建二层ACL 4011，规则中拒绝 802.1p优先级为3的报文（采用自动分配规则号方式，此

时不需要手工具体指定规则号）。

<HUAWEI>system-view

[HUAWEI] acl 4011

[HUAWEI-acl-L2-4011] rule deny 8021p 3

【示例 2】创建二层ACL 4011，定义规则 1，禁止从MAC地址 000d-88f5-97ed发送到MAC地址0011-4301-991e，且802.1p优先级为3的报文通过。

<HUAWEI>system-view

[HUAWEI] acl 4011

[HUAWEI-acl-L2-4011] rule 1 deny 8021p 3 destination-mac 0011-4301-991e ffff-ffff-ffff source-mac 000d-88f5-97ed ffff-ffff-ffff

【示例 3】创建二层ACL 4011，定义规则 1，允许VLAN编号为 2～10的报文通过。VLAN 2～10这个范围对应有掩码为 0xff3，因为只要前面 16位和最后 2位一样时就可以使得VLAN ID范围在 2～10之间。

[HUAWEI] acl 4011

[HUAWEI-acl-L2-4011] rule 1permit vlan-id 2 0xff3

【示例 4】创建二层ACL 4011（采用自动分配规则号方式，此时不需要手工具体指定规则号），规则中允许ARP（类型值为0x0806）报文通过，但拒绝RARP（类型值为0x8035）报文通过。

<HUAWEI>system-view

[HUAWEI] acl 4011

[HUAWEI-acl-L2-4011] rule permit l2-protocol 0x0806

[HUAWEI-acl-L2-4011] rule deny l2-protocol 0x8035

【示例 5】在ACL 4001中采用自动分配ACL规则号方式增加一条规则，匹配目的MAC地址是0000-0000-0001，源MAC地址是0000-0000-0002，二层协议类型值为0x0800的报文。

<HUAWEI>system-view

[HUAWEI] acl 4001

[HUAWEI-acl-L2-4001] rule permit destination-mac 0000-0000-0001 source-mac 0000-0000-0002 l2-protocol 0x0800

9.2.4 配置用户自定义ACL

用户自定义ACL（简称UCL）可以根据用户自定义的规则对数据报文做出相应的处理。它的配置任务也与前面介绍的基本ACL的配置任务是一样的。在具体配置步骤上不同的也只是创建的 ACL类型不同，而且 ACL规则所使用的参数和选项不同。下面同样也仅介绍用户自定义ACL的创建和规则的配置。但**S2700**系列不支持用户自定义**ACL**。

1．用户自定义ACL的创建

在用户自定义ACL的创建中，同样可以创建数字型的或者命名型的ACL。如果创建的是数字型二层ACL，则创建方法与前面介绍的基本ACL的创建方法一样，使用的也是 acl [number] acl-number [match-order {auto | config }] 命令，但其中 acl-number参数的取值范围只能是 5 000～5 999，同样，[match-order {auto | config }] 可选项仅S7700/9300/9700系列交换机支持。

如果创建的是命名型二层ACL，则需要采用 acl nameacl-name {user | acl-number } [match-order { auto | config }] 命令，二选一选项user用来表示所创建的是命名型用户自定义ACL，如果选择 acl-number二选一参数，则其取值范围也是 5 000～5 999，表示创建的是命名型用户自定义ACL。

2．用户自定义ACL规则的配置

二层ACL规则的配置也比基本ACL中的规则配置要复杂许多，可用来匹配的过滤条件比较多，而且也基本上是可同时配置的。具体命令如下：

rule [rule-id] {deny |permit } [[l2-head | ipv4-head | ipv6-head | l4-head] { rule-string rule-mask offset } & <1-8>] [time-range time-name]

说明

S2700/3700/5700EI/5700SI/5700LI/5700S-LI 系列交换机不支持&<1-8>参数，S2700/3700/5700SI/5700LI/5700S-LI系列交换机不支持ipv6-head多选一可选项。

命令中的参数和选项如表9-8所示。

表9-8 用户自定义ACL规则配置命令中的参数和选项说明

参数	说明
rule-id	可选参数，用来指定用户自定义 ACL 规则的编号，取值范围为 0～4 294 967 294 的整数。其他的说明参见 9.2.1 节表 9-4 中的第 6 步
deny	二选一选项，表示拒绝符合条件的报文
permit	二选一选项，表示允许符合条件的报文
l2-head | **ipv4-head** | **ipv6-head** | **l4-head**	可选项，指定 ACL 规则匹配报文时开始偏移的位置（具体的偏移量由后面的 *offset*）参数决定，4 个多选一选项说明如下。 （1）**l2-head**：指定从报文的二层头部开始偏移 （2）**ipv4-head**：指定从 IPv4 头部开始偏移 （3）**ipv6-head**：指定从 IPv6 头部开始偏移 （4）**l4-head**：指定从四层协议头部开始偏移 用户自定义 ACL 就是根据这些报头中的字符串进行匹配的，所以使用用户自定义 ACL 时，一定要对各种协议报头格式非常清楚，一般来说比较难
rule-string	指定 ACL 规则中用于与报文进行匹配的字符串，为 3～10 位十六进制字符，最大支持 4 个字节。命令**每次固定匹配 4 个字节的内容**，当配置的 *rule-string* 参数长度不满 4 个字节时，在前面补 0 凑足 4 个字节进行匹配
rule-mask	指定用于规则字符串匹配的掩码，为 3～10 的十六进制数，最大支持 4 个字节，且必须与 *rule-string* 参数值的长度相同，用于和数据包进行逻辑"与"操作。当用户定义的规则字符串对应的掩码为"1"时，ACL 对该位字符进行匹配；当用户定义的规则字符串对应的掩码为"0"时，ACL 不对该位字符进行匹配（**这与在本章前面介绍的通配符掩码是相反的**）
offset	指定 ACL 规则匹配报文的偏移值，即以 **l2-head** | **ipv4-head** | **ipv6-head** | **l4-head** 选项指定的偏移报头开始从第几个字节进行"与"操作。*rule-mask offset* 参数对共同作用，将从报文提取出来的字符串和用户定义的 *rule-string* 比较，找到匹配的报文，然后进行相应的处理
&<1-8>	指前面的{ *rule-string rule-mask offset* }参数对最多可以有 8 个，以便指定多段用于规则匹配的报头字符串
time-range *time-name*	可选参数，指定使用一个 ACL 规则生效的时间段，参数 *time-name* 用来指定所使用的时间段名称，是 1～32 个字符

【示例 】在编号为 5001 的 ACL 中采用自动分配规则号方式增加一条规则，从二层报文头开始偏移14个字节匹配4个字节的字符串，字符串内容为0180C200。

<HUAWEI>system-view

[HUAWEI] acl 5001

[HUAWEI-acl-user-5001] rule permit l2-head 0x0180C200 0xFFFFFFFF 14

9.2.5 ACL管理

在完成以上 ACL 配置后，或者在出现 ACL 应用问题时可在任意视图下通过以下display命令查看相关

ACL配置信息，验证ACL相关配置结果，或者在用户视图下通过以下reset命令清除ACL相关统计信息。

（1）使用 display acl { acl-number |name acl-name | all }命令查看ACL的相关配置信息。

（2）使用display time-range { all | time-name }命令查看当前时间段的配置和状态。

（3）使用display acl resource [slot slot-id] 命令查看设备上ACL的资源分配信息。可选参数slot-id用来指定要查看资源分配信息的槽位信息：在非堆叠情况下，只能是0；在堆叠情况下，表示堆叠ID。如果不指定此参数，则显示所有堆叠交换机的ACL资源使用情况。

（4）使用 reset acl counter {name acl-name | acl-number | all }命令清除ACL的统计信息。当用户需要清除系统中 ACL 的统计信息，或者当用户需要精确了解某段时间内ACL的统计信息时，可以执行该命令清除系统中ACL的历史统计信息。但执行本命令清除系统中ACL的统计信息时系统不会产生提示信息，所以在执行该命令前，请务必确认清楚。

9.3 基于ACL的简化流策略

上节介绍的是ACL的配置方法，但这些ACL只有在具体位置或功能上得到应用后才会生效，本节所介绍的其实就是ACL的其中一种应用方法（路由信息过滤和策略路由等方面的应用不包括在本节介绍的范围之中）。ACL 可以在交换机上全局，VLAN 或具体接口上应用，但在华为 S 系列交换机中不是直接应用的，是通过一种称之为“基于ACL的简化流策略”来进行的。如可把ACL应用于简化流策略中的多种流行为应用中，如报文过滤、流量监管、流量镜像、流量重定向、报文重标记等。但只有 S2700/3700/5700/6700系列部分机型支持该功能。

说明

本节的内容涉及第10章将要介绍的QoS功能，如QoS流策略中的流量监管、 流量镜像、流量重定向、报文重标记、流量统计等行为，大家可以从第 10章了解相关知识。

9.3.1 基于ACL的简化流策略概述

“基于ACL的简化流策略”是指通过将报文信息与ACL规则进行匹配，为符合ACL规则的报文提供相同的QoS服务，实现对不同类型业务的差分服务。在用户希望对进入网络的流量进行控制时，可以配置根据报文的源IP地址、分片标记、目的IP地址、源端口号、源MAC地址、目的MAC地址等信息的ACL规则对报文进行匹配，进而配置基于ACL的简化流策略实现对匹配ACL规则的报文的过滤监管、重标记、统计、流镜像或重定向。

与第10章将要介绍的QoS中基于流分类的流策略相比，基于ACL的简化流策略不需要单独创建流分类、流行为或流策略，直接通过一条命令 把所采用的基于ACL的流分类和对应的流行为进行关联，达到最终的QoS流策略的目的，配置更为简洁。但是由于仅基于ACL规则对报文进行匹配，因此，匹配规则没有基于QoS流分类的流策略那样丰富，如在ACL中不能配置基于内/外层VLAN标签、优先级映射、报文颜色等规则。有关QoS方面的详细内容将在第10章介绍。

华为 S 系列交换机中的“基于 ACL 的简化流策略”包括报文过滤、流量监管、流量镜像、流量统计、流量重定向、报文重标记、流量统计等几个方面，也可算是 ACL在QoS流策略中的几种主要应用。这些基于ACL的简化流策略都可以应用在交换机全局、具体VLAN或具体交换机接口上。下面分别介绍ACL在以上几种简化流策略中的应用配置方法。

9.3.2 配置基于ACL的报文过滤

通过配置基于ACL的报文过滤，可对匹配ACL规则报文进行禁止/允许动作，进而实现对网络流量的控制（**S2700/3700**系列仅可在入方向应用，**S5700/6700**系列既可以入方向应用，又可在出方向上应用 ）。在基于ACL的简化流策略的报文过滤应用中，用户可以根据以下原则选择使用traffic-filter或traffic-secure命令配置报文过滤。

（1）如果traffic-filter或traffic-secure命令关联的ACL没有同时被其他基于ACL的简化流策略所关联（即两个简化流策略所关联的不是同一个 ACL），且报文不会同时匹配本命令中调用的ACL规则和其他简化流策略关联的ACL规则（即报文不同时匹配 两个简化流策略所关联的ACL规则）时，这两个命令可以任选其一。

（2）如果traffic-filter或traffic-secure命令关联的ACL同时被其他基于ACL的简化流策略所关联（即两个简化流策略所关联的是同一个 ACL），或者报文同时匹配了本命令中调用的ACL规则和其他简化流策略关联的ACL规则（即报文同时匹配 两个简化流策略所关联的ACL规则）时，traffic-filter和traffic-secure的区别如下。

① 当 traffic-secure命令和其他基于ACL的简化流策略同时配置，且ACL规则中的动作为deny时，则仅traffic-secure、traffic-mirror（用来配置根据ACL进行流镜像）和traffic-statistics（用来配置根据ACL进行流量统计）命令的配置将生效（言外之意就是traffic-filter命令的配置不生效），报文被过滤。

② 当 traffic-secure和其他基于ACL的简化流策略同时配置，且ACL规则中的动作为permit时，traffic-secure命令和其他基于ACL的简化流策略均生效。

③ 当 traffic-filter和其他基于ACL的简化流策略同时配置，且ACL规则中的动作为deny时，则仅traffic-filter、traffic-mirror和traffic-statistics命令的配置生效（言外之意就是traffic-secure命令的配置不生效），报文被过滤。

④ 当 traffic-filter和其他基于ACL的简化流策略同时配置，且ACL规则中的动作为permit时，先配置的简化流策略生效。

说明

traffic-secure 命令的优先级高于其他简化 ACL 配置命令，即如果 traffic-secure 命令和其他简化ACL配置命令的配置相冲突时，最终以traffic-secure命令配置为准。

1．在全局或VLAN上应用基于ACL的报文过滤

ACL可在全局或VLAN上应用，配置报文过滤功能，但每个调用的**ACL**仅可匹配一个**ACL**规则，若ACL中包括有许多规则，则必须指出所要应用的具体ACL规则编号。此时需要在系统视图下据实际需要选择以下对应的命令进行配置。

（1）入方向报文过滤

① 在除 S2700-52P-EI/2700-52P-PWR-EI系列之外的其他 S2700EI系列交换机上，执行 traffic-filter [vlan vlan-id] inbound acl {bas-acl | adv-acl } [rule rule-id] 命令对匹配单个ACL规则的入方向报文进行过滤。

② 在 S2752P-EI/2752P-PWR-EI/3700SI/3700EI 系列交换机上，执行 traffic-filter [vlan vlan-id] inbound acl { bas-acl |adv-acl | user-acl } [rule rule-id] 命令对匹配单个ACL规则的入方向报文进行过滤。

③ 在 S5700/6700 系列交换机上，执行 traffic-filter [vlan vlan-id] inbound acl { [ipv6] {bas-acl | adv-acl |nameacl-name } | l2-acl | user-acl } [rule rule-id] 命令对匹配单个ACL规则的入方向报文进行过滤。

④ 在除 S7700/9300/9300E/9700 系列交换机外的其他所有 S 系列交换机上，执行traffic-secure [vlan vlan-id] inboundacl {bas-acl |adv-acl | l2–acl |nameacl-name } [rule rule-id] 命令，对匹配单个ACL规则的入方向报文进行过滤。

⑤ 在除 S7700/9300/9300E/9700 系列交换机外的其他所有 S 系列交换机上执行traffic-secure [vlan vlan-id

] inbound acl { l2–acl |name acl-name }[rule rule-id] acl { ba s-acl | adv-acl |nameacl-name } [rule rule-id] 命令，对同时匹配二层ACL和三层ACL规则的入方向报文进行过滤。

（2）出方向报文过滤

在 S5700/6700系列交换机上（其他系列不支持），执行 traffic-filter [vlan vlan-id] outbound acl { [ipv6] {bas-acl | adv-acl |nameacl-name } | l2-acl } [rulerule-id] 命令对匹配单个ACL规则的出方向报文进行过滤。

（3）入或出方向报文过滤

在 S5700/6700系列交换机上（其他系列不支持），执行 traffic-filter [vlan vlan-id] { inbound | outbound } acl { l2-acl |name acl-name } [rule rule-id] acl {bas-acl | adv-acl |nameacl-name } [rule rule-id] 或 traffic-filter [vlan vlan-id] { inbound | outbound } acl {bas-acl |adv-acl |nameacl-name } [rule rule-id] acl { l2-acl |nameacl-name } [rule rule-id] 命令对同时匹配二层ACL和三层ACL规则的入或出方向报文进行过滤。

以上各个traffic-filter和traffic-secure命令中的参数和选项说明如表9-9所示。

表9-9 traffic-filter和 traffic-secure命令参数和选项说明

参数	说明
vlan *vlan-id*	可选参数，指定在特定 VLAN 上应用基于 ACL 的报文过滤，参数 *vlan-id* 用来指定特定 VLAN 的 VLAN ID，取值范围为 1～4 094 的整数
inbound	指定在入方向上应用报文过滤
outbound	指定在出方向上应用报文过滤，**但基于用户自定义 ACL 的报文过滤不能在出方向上应用**
acl	指定基于 IPv4 ACL 对报文进行过滤
ipv6	可选项，指定基于 IPv6 ACL 对报文进行过滤
bas-acl	多选一参数，指定采用指定编号的基于基本 ACL（可以是基本 ACL 或基本 ACL6）进行报文过滤，取值范围为 2 000～2 999 的整数
adv-acl	多选一参数，指定采用指定编号的基于高级 ACL（可以是高级 ACL 或高级 ACL6）进行报文过滤，取值范围为 3 000～3 999 的整数
l2-acl	多选一参数，指定采用指定编号的基于二层 ACL 进行报文过滤，取值范围为 4 000～4 999 的整数
user-acl	多选一参数，指定采用指定编号的基于用户自定义 ACL 进行报文过滤，取值范围为 5 000～5 999 的整数
name *acl-name*	多选一参数，指定采用指定名称的基于命名型 ACL 进行报文过滤，为 1～32 个字符，不支持空格，区分大小写，且要以英文字母 a～z 或 A～Z 开始
rule *rule-id*	可选参数，指定基于 ACL 中特定规则进行报文过滤。参数 *rule-id* 用来指定对应的规则编号，对于 IPv4 的 ACL，取值范围是 0～4 294 967 294；对于 IPv6 的 ACL，取值范围是 0～2 047

【示例 1】在交换机VLAN 100中（即将在所有加入了VLAN 100的接口上应用）应用基于ACL的报文过滤，仅允许源IP地址为192.168.0.2的主机的IP报文通过，丢弃其他报文。这里需要同时过滤报文协议类型（IP 协议）和源 IP 地址信息，所以需要采用高级ACL。如果不限制报文的协议类型，则可直接用基本ACL来配置。

```
<HUAWEI>system-view
[HUAWEI] vlan 100
[HUAWEI-Vlan100] quit
[HUAWEI] acl name test 3000
[HUAWEI-acl-adv-test] rule 5permit ip source 192.168.0.2 0    #---这里的通配符掩码为0，表示为主机 IP地址
[HUAWEI-acl-adv-test] rule 10deny ip source any
[HUAWEI-acl-adv-test] quit
[HUAWEI] traffic-filter vlan 100 inbound acl name test
```

【示例 2】在交换机全局（所有接口、所有VLAN中）上配置基于ACL的报文过滤功能，将源IP地址为

192.168.0.2的IP报文丢弃。

<HUAWEI>system-view

[HUAWEI] acl 3000

[HUAWEI-acl-adv-3000] rule 5deny ip source 192.168.0.2 0

[HUAWEI-acl-adv-3000] quit

[HUAWEI] traffic-secure inbound acl 3000

2．在端口上应用基于ACL的报文过滤

ACL还可在具体以太网端口上应用基于ACL的报文过滤功能。此时可根据实际需要在具体以太网接口视图下选择使用以下命令进行配置。

（1）入方向报文过滤

① 在除 S2700-52P-EI/2700-52P-PWR-EI系列之外的其他 S2700EI系列交换机上，执行 traffic-filter inbound acl {bas-acl | adv-acl } [rule rule-id] 命令对匹配单个ACL规则的入方向报文进行过滤。

② 在 S2752P-EI/2752P-PWR-EI/3700S/3700EI 系列交换机上，执行 traffic-filter inbound acl {bas-acl | adv-acl | user-acl } [rule rule-id] 命令对匹配单个 ACL规则的入方向报文进行过滤。

③ 在S5700/6700系列交换机上，执行 traffic-filter inbound acl { [ipv6] {bas-acl |adv-acl |name acl-name } | l2-acl | user-acl } [rule rule-id] 命令对匹配单个ACL规则的入方向报文进行过滤。

④ 或者在除 S7700/9300/9700 系列交换机外的其他所有 S 系列交换机上，执行traffic-secure inbound acl { bas-acl | adv-acl | l2–acl |nameacl-name } [rule rule-id] 命令对匹配单个ACL规则的入方向报文进行过滤。

⑤ 在除 S7700/9300/9700 系列交换机外的其他所有 S 系列交换机上，执行traffic-secure inbound acl { l2–acl |name acl-name }

[rule rule-id] acl {bas-acl | adv-acl |name acl-name } [rule rule-id] 命令对同时匹配二层ACL和三层ACL规则的入方向报文进行过滤。

（2）出方向报文过滤

在 S5700/6700 系列交换机上（其他系列不支持），执行 traffic-filter outbound acl { [ipv6] {bas-acl | adv-acl |nameacl-name } | l2-acl } [rule rule-id] 命令对匹配单个 ACL规则的出方向报文进行过滤。

（3）入或出方向报文过滤

在 S5700/6700 系列交换机上（其他系列不支持），执行 traffic-filter { inbound |outbound }acl { l2-acl |nameacl-name } [rule rule-id] acl {bas-acl |adv-acl|nameacl-name } [rule rule-id] 或 traffic-filter { inbound |outbound }acl {bas-acl |adv-acl |nameacl-name } [rule rule-id] acl { l2-acl |nameacl-name } [rule rule-id] 命令对同时匹配二层ACL和三层ACL规则的入或出方向报文进行过滤。

以上traffic-filter和traffic-secure命令的参数和选项说明参见表9-12。

【示例 3】在交换机 GE0/0/1 接口上应用基于 ACL 的报文过滤功能，允许源 IP 为192.168.0.2的主机IP报文通过。

<HUAWEI>system-view

[HUAWEI] acl 3000

[HUAWEI-acl-adv-3000] rule 5permit ip source 192.168.0.2 0

[HUAWEI-acl-adv-3000] quit

[HUAWEI] interface gigabitethernet 0/0/1

[HUAWEI-GigabitEthernet0/0/1] traffic-filter inbound acl 3000

【示例 4】在交换机 GE0/0/1 接口上应用基于 ACL 的报文过滤功能，将源 IP 为192.168.0.2的主机IP报

文丢弃。

```
<HUAWEI>system-view
[HUAWEI] acl 3000
[HUAWEI-acl-adv-3000] rule 5deny ip source 192.168.0.2 0
[HUAWEI-acl-adv-3000] quit
[HUAWEI] interface gigabitethernet 0/0/1
[HUAWEI-GigabitEthernet0/0/1] traffic-secure inbound acl 3000
```

9.3.3 配置基于ACL的流量监管

通过配置基于ACL的流量监管，对匹配ACL规则的报文进行限速，并配置对不同颜色报文采取的动作（**S2700/3700**系列交换机中仅可在入方向上应用**ACL**，**S5700/6700**系列交换机既可在入方向应用，又可以在出方向上应用**ACL** ）。同样，在配置基于ACL的报文过滤之前，需要配置好相应的ACL规则，但每个调用的**ACL**仅可匹配一个**ACL**规则，当ACL中包括许多规则时，则必须指出所要应用的具体ACL规则编号。

1．在全局或VLAN上应用基于ACL的流量监管

在全局或VLAN上应用基于ACL的流量监管的配置需要在系统视图下根据不同S系列交换机选择以下不同命令进行。

（1）在S2700-52P-EI/2700-52P-PWR-EI/2710SI/3700SI/3700EI系列交换机上

执行 traffic-limit [vlan vlan-id] inbound acl { { [ipv6] { bas-acl | adv-acl |name acl-name } } | l2-acl | user-acl } [rule rule-id] cir cir-value [pirpir-value] [cbs cbs-value pbspbs-value] [green {drop |pass [remark-8021p8021p-value |remark-dscpdscp-value] }] [yellow {drop |pass [remark-8021p8021p-value |remark-dscpdscp-value] }] [red {drop |pass [remark-8021p8021p-value | remark-dscpdscp-value] }] 命令对匹配单个ACL规则的入方向报文进行流量监管。

执行 traffic-limit [vlan vlan-id] inbound acl {l2-acl |nameacl-name } [rule rule-id] acl {bas-acl |adv-acl |nameacl-name } [rule rule-id] cir cir-value [pirpir-value] [cbs cbs- valuepbspbs-value] [green {drop |pass [remark-8021p8021p-value |remark-dscpdscp- value] }] [yellow {drop |pass [remark-8021p8021p-value | remark-dscpdscp-value] }] [red { drop |pass [remark-8021p8021p-value | remark-dscpdscp-value] }] 命令对同时匹配二层ACL和三层ACL的入方向报文进行流量监管。

（2）在除S2700-52P-EI和S2700-52P-PWR-EI系列之外的其他S2700EI系列交换机上

执行 traffic-limit [vlan vlan-id] inbound acl { { [ipv6] { bas-acl | adv-acl |name acl- name } } | l2-acl } [rule rule-id] cir cir-value [cbs cbs-value] 命令对匹配单个ACL规则的入方向报文进行流量监管。

执行 traffic-limit [vlan vlan-id] inbound acl { l2-acl |nameacl-name } [rule rule-id] acl {bas-acl | adv-acl |name acl-name } [rule rule-id] cir cir-value [cbscbs-value] 命令对同时匹配二层ACL和三层ACL的入方向报文进行流量监管。

（3）在S5700LI/5700SI/5700S-LI系列交换机上

执行 traffic-limit [vlan vlan-id] inbound acl { [ipv6] {bas-acl | adv-acl |name acl- name } | l2-acl | user-acl } [rule rule-id] cir cir-value [pirpir-value] [cbs cbs-valuepbs pbs-value] [greenpass] [yellow {drop|pass [remark-8021p8021p-value | remark-dscp dscp-value] }] [red {drop |pass [remark-8021p8021p-value |remark-dscpdscp-value] }] 命令对匹配单个ACL规则的入方向报文进行流量监管。

执行 traffic-limit [vlan vlan-id] outbound acl { [ipv6] { bas-acl | adv-acl |name acl- name } | l2-acl } [rulerule-id] } cir cir-value [pirpir-value] [cbs cbs-valuepbspbs-value] [greenpass] [yellowpass] [red {drop |pass

}] 命令对匹配单个ACL规则的出方向报文进行流量监管。

执行 traffic-limit [vlan vlan-id] inbound acl { l2-acl |nameacl-name } [rule rule-id] acl {bas-acl |adv-acl|nameacl-name } [rule rule-id] cir cir-value [pirpir-value] [cbs cbs- valuepbspbs-value] [greenpass] [yellow {drop |pass [remark-8021p8021p-value |remark- dscpdscp-value] }] [red {drop |pass [remark-8021p8021p-value |remark-dscpdscp-value] }] 命令对同时匹配二层ACL和三层ACL的入方向报文进行流量监管。

执行 traffic-limit [vlan vlan-id] outboundacl { l2-acl |nameacl-name } [rule rule-id] acl {bas-acl |adv-acl |nameacl-name } [rule rule-id] cir cir-value [pirpir-value] [cbs cbs- valuepbs pbs-value] [greenpass] [yellowpass] [red {drop |pass }] 命令对同时匹配二层ACL和三层ACL的出方向报文进行流量监管。

（4）其他S5700系列和S6700系列交换机上

执行 traffic-limit [vlan vlan-id] outbound acl { [ipv6] { bas-acl | adv-acl |name acl- name } | l2-acl } [rulerule-id] cir cir-value [pirpir-value] [cbs cbs-valuepbs pbs-value] [[green {drop |pass [remark-8021p8021p-value | remark-dscpdscp-value] }] [yellow {drop |pass [remark-8021p 8021p-value | remark-dscpdscp-value] }] [red {drop |pass [remark- 8021p8021p-value | remark-dscpdscp-value] }]] 命令对匹配单个ACL规则的出方向报文进行流量监管。

执行 traffic-limit [vlan vlan-id] inbound acl { l2-acl |nameacl-name } [rule rule-id] acl {bas-acl | adv-acl|nameacl-name } [rule rule-id] cir cir-value [pirpir-value] [cbs cbs-valuepbs pbs-value] [green {drop |pass [remark-8021p8021p-value | remark-dscp dscp-value] }] [yellow {drop |pass [remark-8021p8021p-value |remark-dscpdscp-value] }] [red { drop |pass [remark-8021p8021p-value | remark-dscpdscp-value] }] 命令对同时匹配二层ACL和三层ACL的入方向报文进行流量监管。

执行 traffic-limit [vlan vlan-id] outboundacl { l2-acl |nameacl-name } [rule rule-id] acl {bas-acl | adv-acl |name acl-name } [rule rule-id] cir cir-value [pirpir-value] [cbs cbs-valuepbs pbs-value] [[green {drop|pass [remark-8021p8021p-value | remark-dscp dscp-value] }] [yellow {drop |pass [remark-8021p8021p-value |remark-dscpdscp-value] }] [red {drop |pass [remark-8021p8021p-value | remark-dscpdscp-value] }]] 命令对同时匹配二层和三层ACL的出方向报文进行流量监管。

说明

报文的颜色可以在流量监管中定义。

（1）报文的突发尺寸<cbs-value时，报文被标记为绿色。

（2）cbs-value≤报文的突发尺寸<pbs-value时，报文被标记为黄色。

（3）报文的突发尺寸≥pbs-value时，报文被标记为红色。

缺省情况下，绿色、黄色报文被允许通过，红色报文被丢弃。

从以上命令可以看出，各命令主要是所支持的参数和选项不同，这些参数和选项的说明如表9-10所示。

表9-10 traffic-limit命令参数和选项说明

参数	说明
vlan *vlan-id*	可选参数，指定要应用简化流策略的 VLAN 的编号，取值范围为 1～4 094 的整数。如果不指定本参数，则将在交换机上全局应用
inbound	指定对入方向报文进行流量监管
outbound	指定对出方向报文进行流量监管
acl	指定基于 IPv4 ACL 对报文进行流量监管
ipv6	可选项，指定基于 IPv6 ACL 对报文进行流量监管
bas-acl	多选一参数，指定基于基本 ACL 对报文进行流量监管的 ACL 编号，取值范围是 2 000～2 999 的整数
adv-acl	多选一参数，指定基于高级 ACL 对报文进行流量监管的 ACL 编号，取值范围是 3 000～3 999 的整数
l2-acl	多选一参数，指定基于二层 ACL 对报文进行流量监管的 ACL 编号，取值范围是 4 000～4 999 的整数
user-acl	多选一参数，指定基于用户自定义 ACL 对报文进行流量监管的 ACL 编号，取值范围是 5 000～5 999 的整数
name *acl-name*	多选一参数，指定基于命名型 ACL 对报文进行过滤。其中，*acl-name* 表示 ACL 的名称，为 1～32 个字符，不支持空格，区分大小写，且要以英文字母 a～z 或 A～Z 开始
rule *rule-id*	指定基于 ACL 中特定规则进行报文流量监管的规则编号。对于 IPv4 ACL，取值范围是 0～4 294 967 294 的整数；对于 IPv6 ACL，取值范围是 0～2 047 的整数。如果 ACL 中仅一条规则，则不用指定本参数

（续表）

参数	说明
cir *cir-value*	可选参数，指定承诺信息速率，即保证能够通过的平均速率，单位是 kbit/s。对于 S5700S-LI/5700LI/5700HI/5710EI/6700 系列交换机，取值范围为 8～10 000 000 的整数；对于 S2700/3700/5700SI/5700EI 系列交换机，取值范围为 64～10 000 000 的整数
pir *pir-value*	可选参数，指定峰值信息速率，即能够通过的最大速率，单位是 kbit/s。对于 S5700S-LI/5700LI/5700HI/5710EI/6700 系列交换机，取值范围为 8～10 000 000 的整数；对于 S2700/3700/5700SI/5700EI 系列交换机，取值范围为 64～10 000 000 的整数
cbs *cbs-value*	可选参数，指定承诺突发尺寸，即瞬间能够通过的承诺突发流量，单位是 byte。S3700/5700/6700 系列交换机的取值范围为 4 000～4 294 967 295 的整数，S2700 系列交换机的取值范围为 8 192～4 294 967 295 的整数。其缺省值与配置的 *cir-value* 有关，对于 S3700/5700/6700 系列交换机，如果 *cir-value**125≤4 000kbit/s，则 *cbs-value* 的缺省值为 4000 bytes。如果 *cir-value**125>4 000kbit/s，则 *cbs-value* 的缺省值等于 *cir-value* 的 125 倍；对于 S2700 系列交换机，如果 *cir-value*≤4096kbit/s，则 *cbs-value* 的缺省值为 4 096byte，如果 *cir-value*>4096kbit/s，则 *cbs-value* 的缺省值等于 *cir-value* 的值
pbs *pbs-value*	可选参数，指定峰值突发尺寸，即瞬间能够通过的峰值突发流量，单位是 byte。取值范围是 4 000～4 294 967 295。缺省值与配置的 *pir-value* 有关，对于 S3700 系列交换机，如果 *pir-value*≤4096kbit/s，则 *pbs-value* 的缺省值为 4096byte，如果 *pir-value*>4096kbit/s，则 *pbs-value* 的缺省值等于 *pir-value* 的值；对于 S5700/6700 系列交换机，如果 *pir-value**125≤4000kbit/s，则 *pbs-value* 的缺省值为 4 000 bytes，如果 *pir-value**125>4 000kbit/s，则 *pbs-value* 的缺省值等于 *pir-value* 的 125 倍
green	可选项，指定对绿色报文进行监管。缺省情况下，绿色报文被允许通过
yellow	可选项，指定对黄色报文进行监管。缺省情况下，黄色报文被允许通过
red	可选项，指定对红色报文进行监管。缺省情况下，红色报文被丢弃
remark *8021p-value*	可选参数，指定重标记报文的 8021p 优先级，取值范围为 0～7 的整数
remark *dscp-value*	可选参数，指定重标记报文的 DSCP 优先级，取值范围为 0～63 的整数
drop	二选一选项，指定丢弃报文
pass	二选一选项，指定允许报文通过

【示例 1】在VLAN100的入方向，配置基于ACL 3000的流量监管功能，其中承诺信息速率为 10 000kbit/s，允许绿色和黄色报文通过，丢弃红色报文。

<HUAWEI>system-view

[HUAWEI] traffic-limit vlan 100 inbound acl 3000 cir 10000 green pass yellow pass red drop

2．在端口上应用基于ACL的流量监管

在端口上应用基于 ACL 的流量监管的配置也要根据不同 S 系列交换机选择对应的以下配置命令在具体的接口视图下进行。

（1）在S2700-52P-EI/2700-52P-PWR-EI/2710SI/3700SI/3700EI系列交换机上

执行 traffic-limit inbound acl { { [ipv6] {bas-acl | adv-acl |nameacl-name } } | l2-acl | user-acl } [rule rule-id] cir cir-value [pir pir-value] [cbs cbs-valuepbspbs-value] [green {drop |pass [remark-8021p8021p-value | remark-dscpdscp-value] }] [yellow {drop |pass [remark-8021p 8021p-value | remark-dscpdscp-value] }] [red {drop |pass [remark-8021p 8021p-value |remark-dscpdscp-value] }] 命令对匹配单个ACL规则的入方向报文进行流量监管。

执行 traffic-limit inbound acl { l2-acl |nameacl-name } [rule rule-id] acl { bas-acl | adv-acl |name acl-name } [rule rule-id] cir cir-value [pirpir-value] [cbs cbs-valuepbs pbs-value] [green {drop |pass [remark-8021p8021p-value | remark-dscpdscp-value] }] [yellow {drop |pass [remark-8021p8021p-value |remark-dscpdscp-value] }] [red {drop |pass [remark-8021p8021p-value | remark-dscpdscp-value] }] 命令对同时匹配二层ACL和三层ACL规则的入方向报文进行流量监管。

（2）在除S2700-52P-EI和S2700-52P-PWR-EI系列之外的其他S2700EI系列交换机上

执行 traffic-limit inboundacl { { [ipv6] {bas-acl|adv-acl |nameacl-name } } | l2-acl } [rule rule-id] cir cir-value [cbs cbs-value] 命令对匹配单个 ACL 规则的入方向报文进行流量监管。

执行 traffic-limit inboundacl { l2-acl |nameacl-name } [rule rule-id] acl { bas-acl |adv-acl |name acl-name } [rule rule-id] cir cir-value [cbs cbs-value] 命令对同时匹配二层ACL和三层ACL规则的入方向报文进行流量监管。

（3）在S5700LI/5700SI/5700S-LI系列交换机上

执行 traffic-limit inbound acl { [ipv6] { bas-acl | adv-acl |nameacl-name } | l2-acl |user-acl } [rule rule-id] cir cir-value [pir pir-value] [cbs cbs-valuepbs pbs-value] [green pass] [yellow {drop |pass [remark-8021p8021p-value | remark-dscpdscp-value] }] [red {drop |pass [remark-8021p8021p-value | remark-dscpdscp-value] }] 命令对匹配单个ACL规则的入方向报文进行流量监管。

执行 traffic-limitoutbound acl { [ipv6] { bas-acl | adv-acl |nameacl-name } | l2-acl } [rule rule-id] } circir-value [pirpir-value] [cbs cbs-valuepbs pbs-value] [greenpass] [yellowpass] [red {drop |pass }] 命令对匹配单个ACL规则的出方向报文进行流量监管。

执行 traffic-limit inboundacl { l2-acl |nameacl-name } [rule rule-id] acl { bas-acl |adv-acl |name acl-name} [rule rule-id] cir cir-value [pirpir-value] [cbs cbs-valuepbs pbs-value] [greenpass] [yellow {drop |pass [remark-8021p8021p-value | remark-dscp dscp-value] }] [red {drop |pass [remark-8021p8021p-value |remark-dscpdscp-value] }] 命令对同时匹配二层ACL和三层ACL的入方向报文进行流量监管。

执行traffic-limitoutbound acl { l2-acl |nameacl-name } [rule rule-id] acl {bas-acl |adv-acl |name acl-name} [rule rule-id] cir cir-value [pirpir-value] [cbs cbs-valuepbs pbs-value] [greenpass] [yellowpass] [red {drop |pass }] 命令对同时匹配二层ACL和三层ACL的出方向报文进行流量监管。

（4）在其他S5700系列和S6700系列交换机上

执行 traffic-limit inbound acl { { [ipv6] {bas-acl | adv-acl |nameacl-name } } | l2-acl | user-acl } [rulerule-id] cir cir-value [pirpir-value] [cbs cbs-valuepbspbs-value] [green {drop |pass [remark-8021p8021p-value | remark-dscpdscp-value] }] [yellow {drop |pass [remark-8021p 8021p-value | remark-dscpdscp-value] }] [red {drop |pass [remark-8021p8021p-value | remark-dscpdscp-value] }] 命令对匹配单个ACL规则的入方向报文进行流量监管。

执行 traffic-limitoutbound acl { [ipv6] { bas-acl | adv-acl |nameacl-name } | l2-acl } [rule rule-id] circir-value [pirpir-value] [cbs cbs-valuepbspbs-value] [[green {drop |pass [remark-8021p8021p-value |remark-dscpdscp-value] }] [yellow {drop |pass [remark-8021p8021p-value | remark-dscpdscp-value] }] [red{drop |pass

[remark- 8021p8021p-value | remark-dscpdscp-value] }]] 命令对匹配单个ACL规则的出方向报文进行流量监管。

执行 traffic-limit inboundacl { l2-acl |nameacl-name } [rule rule-id] acl { bas-acl |adv-acl |name acl-name} [rule rule-id] cir cir-value [pirpir-value] [cbs cbs-valuepbs pbs-value] [green {drop |pass [remark-8021p8021p-value | remark-dscpdscp-value] }] [yellow {drop |pass [remark-8021p8021p-value | remark-dscpdscp-value] }] [red { drop |pass [remark-8021p 8021p-value | remark-dscpdscp-value] }] 命令对同时匹配二层ACL和三层ACL规则的入方向报文进行流量监管。

执行traffic-limitoutbound acl { l2-acl |nameacl-name } [rule rule-id] acl {bas-acl |adv-acl |name acl-name} [rule rule-id] cir cir-value [pirpir-value] [cbs cbs-valuepbs pbs-value] [[green {drop |pass [remark-8021p8021p-value |remark-dscpdscp-value] }] [yellow {drop |pass [remark-8021p8021p-value | remark-dscpdscp-value] }] [red {drop |pass [remark-8021p 8021p-value | remark-dscpdscp-value] }]] 命令对同时匹配二层ACL和三层ACL规则的出方向报文进行流量监管。

从以上命令可以看出，各命令主要是所支持的参数和选项不同，这些参数和选项的说明参见表9-12。

【示例 2】在GE0/0/1接口入方向配置基于ACL的流量监管功能。配置匹配ACL 3000的报文的承诺信息速率为 10 000kbit/s，允许绿色、黄色、红色报文通过，重标记红色报文的DSCP优先级为5。

<HUAWEI > system-view

[HUAWEI] interface gigabitethernet 0/0/1

[HUAWEI-GigabitEthernet0/0/1] traffic-limit inbound acl 3000 cir 10000 green pass yellow pass red pass remark dscp 5

9.3.4 配置基于ACL的流镜像

通过配置基于ACL的流镜像可将匹配ACL规则的报文镜像到指定观察接口，以便于对报文进行分析（均仅可在入方向上应用**ACL** ）。在配置基于ACL的报文过滤之前需要配置好相应的ACL规则，但每个调用的**ACL**仅可匹配一个**ACL**规则，当ACL中包括有许多规则时，必须指出所要应用的具体ACL规则编号。

1．在全局或VLAN上应用基于ACL的流镜像

在全局或VLAN上应用基于ACL的流镜像配置也是在系统视图下根据不同S系列交换机选择以下不同命令进行的。

（1）在S2700-52P-EI/2700-52P-PWR-EI/2710SI/3700SI/3700EI系列交换机上

执行 traffic-mirror [vlan vlan-id] inbound acl{ { [ipv6] {bas-acl | adv-acl |name acl-name } } | l2-acl | user-acl } [rule rule-id] toobserve-porto-index 命令对匹配单个ACL规则的入方向报文进行流镜像。

执行traffic-mirror [vlan vlan-id] inboundacl { l2-acl |nameacl-name } [rule rule-id] acl {bas-acl |adv-acl |nameacl-name } [rule rule-id] toobserve-porto-index [remotevlan-id] 命令对同时匹配二层ACL和三层ACL规则的入方向报文进行流镜像。

（2）在除S2700-52P-EI和S2700-52P-PWR-EI系列之外的其他S2700EI系列交换机上

执行 traffic-mirror [vlan vlan-id] inbound acl { { [ipv6] {bas-acl | adv-acl |name acl-name } } | l2-acl } [rule rule-id] to observe-porto-index命令对匹配单个ACL规则的入方向报文进行流镜像。

（3）在S5700/6700系列交换机上

执行 traffic-mirror [vlan vlan-id] inbound { acl { [ipv6] {bas-acl | adv-acl |name acl-name } | l2-acl |user-acl } } [rule rule-id] to observe-porto-index [remote vlan-id] 命令对匹配单个ACL规则的入方向报文进行流镜像。

根据需要选择以下一个命令对同时匹配二层ACL和三层ACL规则的入方向报文进行流镜像。

（1）traffic-mirror [vlanvlan-id] inboundacl l2-acl [rule rule-id] acl {bas-acl |adv-acl |name acl-name } [rule rule-id] to observe-porto-index [remote vlan-id]

（2）traffic-mirror [vlan vlan-id] inbound acl name acl-name [rule rule-id] acl {bas- acl |adv-acl | l2-acl |nameacl-name } [rule rule-id] to observe-porto-index [remote vlan-id]

以上traffic-mirror命令中的参数和选项说明如表9-11所示。

表9-11 traffic-mirror命令参数和选项说明

参数	说明
vlan *vlan-id*	指定要应用简化流镜像策略的 VLAN 的编号，取值范围为 1～4 094 的整数
inbound	表示对入方向的报文进行流镜像
acl	指定基于 IPv4 ACL 对报文进行流镜像
ipv6	指定基于 IPv6 ACL 对报文进行流镜像
bas-acl	指定基于基本 ACL 对报文进行流镜像的 ACL 编号，取值范围为 2 000～2 999 的整数
adv-acl	指定基于高级 ACL 对报文进行流镜像的 ACL 编号，取值范围为 3 000～3 999 的整数
l2-acl	指定基于二层 ACL 对报文进行流镜像的 ACL 编号，取值范围为 4 000～4 999 的整数
user-acl	指定基于用户自定义 ACL 对报文进行流镜像的 ACL 编号，取值范围为 5 000～5 999 的整数
name *acl-name*	指定基于命名型 ACL 对报文进行过滤。其中，*acl-name* 表示 ACL 的名称，为 1～32 个字符，不支持空格，区分大小写，且要以英文字母 a～z 或 A～Z 开始
rule *rule-id*	指定基于 ACL 中特定规则进行报文流镜像的规则编号。对于 IPv4 ACL，取值范围为 0～4 294 967 294 的整数；对于 IPv6 ACL，取值范围为 0～2 047 的整数
to observe-port *o-index*	指定报文镜像到的全局观察端口的索引号，**要先配置好对应索引号的观察端口。** S2700/3700SI 5700SI/5700LI/5700S-LI/6700 系列交换机仅可取值 1，S5700HI/5710EI 系列交换机可以取值 1 或 2，S3700EI/5700EI 系列交换机的取值范围为 1～4 的整数
remote *vlan-id*	指定观察端口所属的 VLAN，取值范围为 1～4 094 的整数。**仅** S5700EI/5700HI/5710EI/6700 **系列交换机支持此可选参数**

【示例 1】在VLAN100的入方向配置基于ACL的流镜像功能，将匹配ACL 3000的报文镜像到索引为1的观察端口。

< HUAWEI > system-view

[HUAWEI] observe-port 1 interface gigabitethernet 0/0/1

[HUAWEI] traffic-mirror vlan 100 inbound acl 3000 to observe-port 1

2．在端口上应用基于ACL的流镜像

在端口上应用基于 ACL 的流镜像的配置也要根据不同 S 系列交换机选择对应的以下配置命令在具体的接口视图中进行。

（1）在S2700-52P-EI/2700-52P-PWR-EI/2710SI/3700SI/3700EI系列交换机上

执行 traffic-mirror inbound acl { { [ipv6] {bas-acl | adv-acl |nameacl-name } } | l2-acl | user-acl } [rule rule-id] to observe-porto-index命令对匹配单个ACL规则的入方向报文进行流镜像。

执行 traffic-mirror inboundacl { l2-acl |nameacl-name } [rule rule-id] acl {bas-acl |adv-acl |name acl-name} [rule rule-id] to observe-porto-index [remote vlan-id] 命令对同时匹配二层ACL和三层ACL规则的入方向报文进行流镜像。

（2）在除S2700-52P-EI和S2700-52P-PWR-EI系列之外的其他S2700EI系列交换机上

执行 traffic-mirror inbound acl { { [ipv6] { bas-acl | adv-acl |nameacl-name } } | l2-acl } [rule rule-id] toobserve-porto-index命令对匹配单个 ACL规则的入方向报文进行流镜像。

（3）在S5700/6700系列交换机上

执行 traffic-mirror [vlan vlan-id] inbound { acl { [ipv6] {bas-acl | adv-acl |name acl-name } | l2-acl |user-acl } } [rule rule-id] to observe-porto-index [remote vlan-id] 命令对匹配单个ACL规则的入方向报文进行流镜

像。

根据需要选择以下一个命令对同时匹配二层ACL和三层ACL规则的入方向报文进行流镜像。

（1）traffic-mirror [vlanvlan-id] inboundacl l2-acl [rule rule-id] acl {bas-acl |adv-acl |name acl-name } [rule rule-id] to observe-porto-index [remote vlan-id]

（2）traffic-mirror [vlan vlan-id] inboundaclnameacl-name [rule rule-id] acl {bas- acl | adv-acl | l2-acl |name acl-name } [rule rule-id] to observe-port o-index [remote vlan-id]

以上traffic-mirror命令中的参数和选项说明参见表9-13。

【示例 2】在接口GE0/0/1的入方向，配置基于ACL的流镜像功能。配置匹配ACL 3000的报文，镜像到索引为1的观察端口。

< HUAWEI > system-view

[HUAWEI] observe-port 1 interface gigabitethernet 0/0/1

[HUAWEI] interface gigabitethernet 0/0/1

[HUAWEI-GigabitEthernet0/0/1] traffic-mirror inbound acl 3000 to observe-port 1

9.3.5 配置基于ACL的重定向

通过配置基于ACL的重定向，将匹配ACL规则的报文重定向到CPU、指定接口或指定下一跳地址（均仅可在入方向上应用**ACL**）。在配置基于ACL的报文过滤之前，需要配置好相应的ACL规则，但每个调用的**ACL**仅可匹配一个**ACL**规则，当ACL中包 括有许多规则时，则必须指出所要应用的具体 ACL 规则编号。但在 **S2700** 系列中，只有**S2700-52P-EI/2700-52P-PWR-EI/2710SI**系列交换机支持基于**ACL**的重定向功能。

1．在全局或VLAN上应用基于ACL的重定向

在全局或VLAN上应用基于ACL的重定向配置也是在系统视图下根据不同S系列交换机选择以下不同命令进行的。

（1）在S2700-52P-EI/2700-52P-PWR-EI/S2710SI/3700系列交换机上

执行 traffic-redirect [vlan vlan-id] inboundacl { [ipv6] {bas-acl | adv-acl |name acl-name } | l2-acl |user-acl } [rule rule-id] { cpu | interface interface-type interface-number |ip-nexthop ip-nexthop | ipv6-nexthopipv6-nexthop } 命令对匹配单个ACL规则的入方向报文进行重定向。

执行 traffic-redirect [vlan vlan-id] inbound acl { l2-acl |nameacl-name } [rule rule- id] acl { bas-acl | adv-acl |nameacl-name } [rule rule-id] { cpu | interface interface-type interface-number | ip-nexthop ip-nexthop |ipv6-nexthop ipv6-nexthop }命令对同时匹配二层ACL和三层ACL的入方向报文进行重定向。

（2）在S5700SI/5700LI/5700S-LI系列交换机上

执行 traffic-redirect [vlan vlan-id] inboundacl { [ipv6] {bas-acl | adv-acl |name acl-name } | l2-acl |user-acl } } [rule rule-id] { cpu | interface interface-type interface- number }命令对匹配单个ACL规则的入方向报文进行重定向。

执行 traffic-redirect [vlan vlan-id] inbound acl { l2-acl |nameacl-name } [rule rule- id] acl { bas-acl | adv-acl |nameacl-name } [rule rule-id] { cpu | interface interface-type interface-number }命令对同时匹配二层和三层ACL规则的入方向报文进行重定向。

（3）在其他S5700/6700系列交换机上

执行 traffic-redirect [vlan vlan-id] inboundacl { [ipv6] {bas-acl | adv-acl |name acl-name } | l2-acl |user-acl } [rule rule-id] { cpu | interface interface-type interface-number |ip-nexthop ip-nexthop | ipv6-nexthopipv6-nexthop }

命令对匹配单个ACL规则的入方向报文进行重定向。

执行 traffic-redirect [vlan vlan-id] inbound acl { l2-acl |nameacl-name } [rule rule- id] acl { bas-acl | adv-acl |nameacl-name } [rule rule-id] { cpu | interface interface-type interface-number | ip-nexthop ip-nexthop |ipv6-nexthop ipv6-nexthop }命令对同时匹配二层ACL和三层ACL的入方向报文进行重定向。

以上traffic-redirect命令中的参数和选项说明如表9-12所示。

表9-12 traffic-redirect命令参数和选项说明

参数	说明
vlan *vlan-id*	指定要应用基于 ACL 的重定向的 VLAN 的编号，取值范围为 1～4 094 的整数
inbound	指定对入方向的报文进行重定向
acl	指定基于 IPv4 ACL 对报文进行重定向
ipv6	指定基于 IPv6 ACL 对报文进行重定向
bas-acl	指定基于基本 ACL 对报文进行重定向的 ACL 编号，取值范围为 2 000～2 999 的整数

（续表）

参数	说明
adv-acl	指定基于高级 ACL 对报文进行重定向的 ACL 编号，取值范围为 3 000～3 999 的整数
l2-acl	指定基于二层 ACL 对报文进行重定向的 ACL 编号，取值范围为 4 000～4 999 的整数
user-acl	指定基于用户自定义 ACL 对报文进行重定向的 ACL 编号，取值范围为 5 000～5 999 的整数
name *acl-name*	指定基于命名型 ACL 对报文进行过滤。其中，*acl-name* 表示 ACL 的名称。其中，*acl-name* 表示 ACL 的名称，为 1～32 个字符，不支持空格，区分大小写，且要以英文字母 a～z 或 A～Z 开始
rule *rule-id*	指定基于 ACL 中特定规则进行报文重定向。对于 IPv4 ACL，取值范围为 0～4 294 967 294 的整数；对于 IPv6 ACL，取值范围为 0～2 047 的整数
cpu	指定将报文重定向到 CPU
interface *interface-type interface-number*	指定将报文重定向到接口
ip-nexthop *ip-nexthop*	指定将报文重定向到下一跳 IPv4 地址
ipv6-nexthop *ipv6-nexthop*	指定将报文重定向到下一跳 IPv6 地址

【示例 1】在VLAN100的入方向，配置基于ACL的重定向功能。将匹配ACL 3000的报文，重定向到接口GE0/0/1。

<HUAWEI>system-view

[HUAWEI] traffic-redirect vlan 100 inbound acl 3000 interface gigabitethernet 0/0/1

2．在端口上应用基于ACL的重定向

在端口上应用基于 ACL 的重定向的配置也要根据不同 S 系列交换机选择对应的以下配置命令在具体的接口视图下进行。

（1）在S2700-52P-EI/2700-52P-PWR-EI/S2710SI/3700系列交换机上

执 行 traffic-redirect inbound acl { [ipv6] {bas-acl | adv-acl |nameacl-name } | l2-acl |user-acl } [rulerule-id] { cpu | interface interface-type interface-number | ip-nexthop ip-nexthop | ipv6-nexthop ipv6-nexthop }命令对匹配单个 ACL 规则的入方向报文进行重定向。

执行 traffic-redirect inboundacl { l2-acl |nameacl-name } [rule rule-id] acl {bas-acl |adv-acl |name acl-name } [rule rule-id] { cpu | interface interface-type interface-number |ip-nexthop ip-nexthop | ipv6-nexthop ipv6-nexthop }命令对同时匹配二层 ACL 和三层ACL的入方向报文进行重定向。

（2）在S5700SI/5700LI/5700S-LI系列交换机上

执 行 traffic-redirect inbound acl { [ipv6] {bas-acl | adv-acl |nameacl-name } | l2-acl | user-acl } [rulerule-id] { cpu | interface interface-type interface-number }命令对匹配单个ACL规则的入方向报文进行重定向。

执行 traffic-redirect inboundacl { l2-acl |nameacl-name } [rule rule-id] acl {bas-acl |adv-acl |name acl-name } [rule rule-id] { cpu | interface interface-type interface-number } 命令对同时匹配二层和三层ACL规则的入方向报文进行重定向。

（3）在其他S5700系列和S6700系列交换机上

执 行 traffic-redirect inbound acl { [ipv6] {bas-acl | adv-acl |nameacl-name } | l2-acl |user-acl } [rulerule-id] { cpu | interface interface-type interface-number | ip-nexthop ip-nexthop | ipv6-nexthop ipv6-nexthop }命令对匹配单个 ACL 规则的入方向报文进行重定向。

执行 traffic-redirect inboundacl { l2-acl |nameacl-name } [rule rule-id] acl {bas-acl |adv-acl |name acl-name } [rule rule-id] { cpu | interface interface-type interface-number |ip-nexthop ip-nexthop | ipv6-nexthop ipv6-nexthop }命令对同时匹配二层 ACL 和三层ACL的入方向报文进行重定向。

以上traffic-redirect命令中的参数和选项说明参见表9-14。

【示例 2】在Eth0/0/1接口入方向，配置基于接口的报文重定向功能。配置ACL 3000的报文，重定向到Eth0/0/2接口。

```
<HUAWEI>system-view
[HUAWEI] interface ethernet 0/0/1
[HUAWEI-Ethernet0/0/1] traffic-redirect inbound acl 3000 interface ethernet 0/0/2
```

9.3.6 配置基于ACL的重标记

通过配置基于 ACL 的重标记可对匹配指定 ACL 规则的报文重标记其优先级，如VLAN报文中的802.1p、IP报文中的DSCP等 （**S2700/3700**系列交换机中仅可在入方向上应用**ACL)**，**(S5700/6700**系列交换机既可在入方向，又可以在出方向上应用**ACL** ）。在配置基于ACL的报文过滤之前也需要配置相应的ACL规则，但每个调用的ACL仅可匹配一个ACL规则，当ACL中包括有许多规则时，则必须指出所要应用的具体ACL规则编号。

1．在全局或VLAN上应用基于ACL的重标记

在全局或VLAN上应用基于ACL的重标记配置也是在系统视图下根据不同S系列交换机选择以下不同命令进行的。

（1）在S2700-52P-EI/2700-52P-PWR-EI/2710SI/3700SI/3700EI系列交换机上

执行 traffic-remark [vlanvlan-id] inboundacl{ { [ipv6] {bas-acl |adv-acl |name acl-name } } | l2-acl |user-acl } [rule rule-id] {dscp {dscp-name |dscp-value } |8021p 8021p-value |destination-macmac-address | ip-precedence ip-precedence-value |vlan-idvlan-id |local-precedence local-precedence-value }命令对匹配单个 ACL 规则的入方向报文进行重标记。

执行traffic-remark [vlan vlan-id] inboundacl{ l2-acl |nameacl-name } [rule rule-id] acl {bas-acl | adv-acl |name acl-name } [rule rule-id] {dscp {dscp-name | dscp-value } | 8021p8021p-value |destination-macmac-address | ip-precedence ip-precedence-value | vl an-idvlan-id | local-precedence local-precedence-value }命令对同时匹配二层 ACL和三层ACL规则的入方向报文进行重定向。

（2）在除S2700-52P-EI和S2700-52P-PWR-EI系列之外的其他S2700EI系列交换机上

执行 traffic-remark [vlan vlan-id] inbound acl { { [ipv6] {bas-acl | adv-acl |name acl-name } } | l2-acl } [rule

rule-id] {dscp { dscp-name | dscp-value } | 8021p8021p-value |vlan-id vlan-id | local-precedence local-precedence-value }命令对匹配单个 ACL规则的入方向报文进行重标记。

执行 traffic-remark [vlan vlan-id] inbound acl {l2-acl |nameacl-name } [rule rule- id] acl { bas-acl | adv-acl |nameacl-name } [rule rule-id] {dscp { dscp-name| dscp-value } | 8021p8021p-value | vlan-id vlan-id | local-precedence local-precedence-value } 命令同时匹配二层ACL和三层ACL规则的入方向报文进行重定向。

（3）在S5700/6700系列交换机上

执行 traffic-remark [vlan vlan-id] inboundacl { [ipv6] {bas-acl |adv-acl |name acl- name } | l2-acl |user-acl } [rule rule-id] {8021p8021p-value |destination-macmac- address |dscp {dscp-name |dscp-value } | local-precedence local-precedence-value | ip- precedence ip-precedence-value | vlan-id vlan-id }命令匹配单个ACL规则的入方向报文进行重标记。

执行 traffic-remark [vlan vlan-id] outboundacl { [ipv6] { bas-acl | adv-acl |name acl-name } | l2-acl } [rule rule-id] {8021p8021p-value | cvlan-id cvlan-id |dscp {dscp-name |dscp-value } | vlan-id vlan-id }命令对匹配单个ACL规则的出方向报文进行重标记。

执行 traffic-remark [vlan vlan-id] inbound acl { l2-acl |nameacl-name } [rule rule- id] acl {bas-acl |adv-acl |nameacl-name } [rule rule-id] {8021p8021p-value |destination- macmac-address |dscp {dscp-name |dscp-value } | local-precedence local-precedence-value |ip-precedence ip-precedence-value | vlan-id vlan-id }命令对同时匹配二层 ACL 和三层ACL规则的入方向报文进行重定向。

执行 traffic-remark [vlan vlan-id] outboundacl { l2-acl |nameacl-name } [rule rule- id] acl { bas-acl | adv-acl |nameacl-name } [rule rule-id] {8021p8021p-value | cvlan-id cvlan-id |dscp { dscp-name | dscp-value } |vlan-id vlan-id }命令同时匹配二层ACL和三层ACL规则的出方向报文进行重定向。

以上traffic-remark命令中的参数和选项说明如表9-13所示。

表9-13 traffic-remark命令参数和选项说明

参数	说明
vlan *vlan-id*	指定要应用基于 ACL 的重标记的 VLAN 的编号，取值范围为 1～4 094 的整数
inbound	指定对入方向的报文进行重标记
outbound	指定对出方向的报文进行重标记
acl	指定基于 IPv4 ACL 对报文进行重标记
ipv6	指定基于 IPv6 ACL 对报文进行重标记
bas-acl	指定基于基本 ACL 对报文进行重标记的 ACL 编号，取值范围为 2 000～2 999 的整数
adv-acl	指定基于高级 ACL 对报文进行重标记的 ACL 编号，取值范围为 3 000～3 999 的整数
l2-acl	指定基于二层 ACL 对报文进行重标记的 ACL 编号，取值范围为 4 000～4 999 的整数
user-acl	指定基于用户自定义 ACL 对报文进行重标记的 ACL 编号，取值范围为 5 000～5 999 的整数

（续表）

参数	说明
name *acl-name*	指定基于命名型 ACL 对报文进行过滤。其中，*acl-name* 表示 ACL 的名称。其中，*acl-name* 表示 ACL 的名称，为 1～32 个字符，不支持空格，区分大小写，且要以英文字母 a～z 或 A～Z 开始
rule *rule-id*	指定基于 ACL 中特定规则进行报文重标记。对于 IPv4 ACL，取值范围为 0～4 294 967 294 的整数；对于 IPv6 ACL，取值范围为 0～2 047 的整数
8021p *8021p-value*	指定重标记报文的 8021p 优先级，取值范围为 0～7 的整数，值越大优先级越高
cvlan-id	指定重标记 QinQ 报文中的内层 VLAN 标签，取值范围为 1～4 094 的整数。**但 S2700/3700/5700SI/5700LI/5700S-LI 系列交换机不支持重标记 QinQ 报文中的内层 VLAN 标签**
destination-mac *mac-address*	指定重标记报文的目的 MAC 地址，格式为 H-H-H，其中 H 为 1 至 4 位的十六进制数。**但 S2700/3700/5700SI/5700LI/5700S-LI 系列交换机不支持重标记报文的目的 MAC 地址**
dscp { *dscp-name* | *dscp-value* }	指定重标记报文的 DSCP 的服务类型。可以为 DiffServ 编码，整数形式，取值范围是 0～63；也可以为 DSCP 的服务类型名称。它们之间的对应关系为 af11（10）、af12（12）、af13（14）、af21（18）、af22（20）、af23（22）、af31（26）、af32（28）、af33（30）、af41（34）、af42（36）、af43（38）、cs1～cs7（分别对应 8、16、24、32、40、48、56）、default（0）、ef（46）。**但 S2700 系列不支持此参数**
local-precedence *local-precedence-value*	指定重标记报文的本地优先级，取值范围为 0～7 的整数，值越大优先级越高
ip-precedence *ip-precedence-value*	指定重标记报文的 IP 优先级，取值范围为 0～7 的整数，值越大优先级越高，**但 S2700 系列不支持此参数**
vlan *vlan-id*	重标记后的 VLAN 编号，取值范围为 1～4 094 的整数

【示例 1】在VLAN100的入方向，配置基于ACL的重标记功能。将源MAC地址为0-0-1的报文，重标记VLAN ID为 101。

<HUAWEI>system-view

[HUAWEI] acl 4001

[HUAWEI-acl-L2-4001] rule 5permit source-mac 0-0-1

[HUAWEI-acl-L2-4001] quit

[HUAWEI] traffic-remark vlan 100 inbound acl 4001 rule 5 vlan-id 101

2．在接口上应用基于ACL的重标记

在端口上应用基于 ACL 的重标记的配置也要根据不同 S 系列交换机选择对应的以下配置命令在具体的接口视图下进行。

（1）在S2700-52P-EI/2700-52P-PWR-EI/2710SI/3700SI/3700EI系列交换机上

执行 traffic-remark inbound acl { { [ipv6] {bas-acl | adv-acl |nameacl-name } } | l2-acl | user-acl } [rule rule-id] {dscp { dscp-name | dscp-value } | 8021p8021p-value |destination-macmac-address | ip-precedence ip-precedence-value | vlan-id vlan-id | loc al-precedencelocal-precedence-value }命令对匹配单个ACL规则的入方向报文进行重标记。

执行 traffic-remark inboundacl { l2-acl |nameacl-name } [rule rule-id] acl {bas-acl | adv- acl |nameacl-name } [rulerule-id] {dscp {dscp-name |dscp-value } |8021p8021p-value | destination- macmac-address | ip-precedence ip-precedence-value | vlan-id vlan-id | local-precedence l ocal-precedence-value }命令对同时匹配二层ACL和三层ACL规则的入方向报文进行重定向。

（2）在除S2700-52P-EI和S2700-52P-PWR-EI系列之外的其他S2700EI系列交换机上

执行 traffic-remark inboundacl { { [ipv6] {bas-acl |adv-acl |nameacl-name } } | l2-acl } [rule rule-id] {dscp { dscp-name | dscp-value } | 8021p8021p-value |vlan-id vlan-id | local- precedence local-precedence-value }命令对匹配单个ACL规则的入方向报文进行重标记。

执行 traffic-remark inboundacl { l2-acl |nameacl-name } [rule rule-id] acl {bas-acl |adv-acl |nameacl-name } [rule rule-id] {dscp {dscp-name |dscp-value } |8021p8021p-value |vlan-id vlan-id | local-precedence local-precedence-value }命令对同时匹配二层ACL和三层ACL规则的入方向报文进行重定向。

（3）在S5700/6700系列交换机上

执 行 traffic-remark inboundacl { [ipv6] {bas-acl |adv-acl |nameacl-name } | l2-acl |user-acl } [rule rule-id]

{8021p8021p-value |destination-macmac-address |dscp {dscp-name |dscp-value } | local-precedence local-precedence-value | ip-precedence ip- precedence-value |vlan-id vlan-id }命令对匹配单个ACL规则的入方向报文进行重标记。

执行 traffic-remarkoutbound acl { [ipv6] { bas-acl | adv-acl |nameacl-name } | l2-acl } [rule rule-id] {8021p8021p-value |cvlan-idcvlan-id |dscp {dscp-name |dscp-value } | vlan-id vlan-id }命令对匹配单个ACL规则的出方向报文进行重标记。

执行 traffic-remark inboundacl { l2-acl |nameacl-name } [rule rule-id] acl {bas- acl |adv-acl |nameacl-name} [rulerule-id] {8021p8021p-value |destination-macmac- address |dscp {dscp-name |dscp-value } | local-precedence local-precedence-value | ip-precedence ip-precedence- value |vlan-idvlan-id }命令对同时匹配二层ACL和三层ACL规则的入方向报文进行重定向。

执行 traffic-remarkoutbound acl { l2-acl |nameacl-name } [rule rule-id] acl {bas- acl | adv-acl |nameacl-name } [rule rule-id] {8021p8021p-value | cvlan-id cvlan-id |dscp { dscp-name |dscp-value } |vlan-id vlan-id }命令对同时匹配二层ACL和三层ACL规则的出方向报文进行重定向。

【示例 2】在接口GE0/0/1的入方向，配置基于ACL的重标记功能。将源MAC地址为 0-0-1的报文，重标记VLAN ID为 100。

```
< HUAWEI > system-view
[HUAWEI] acl 4001
[HUAWEI- acl-L2-4001] rule 5 permit source-mac 0-0-1
[HUAWEI-acl-L2-4001] quit
[HUAWEI] interface gigabitethernet 0/0/1
[HUAWEI-GigabitEthernet0/0/1] traffic-remark inbound acl 4001 rule 5 vlan-id 100
```

9.3.7 配置基于ACL的流量统计

通过配置基于 ACL 的流量统计可对匹配指定 ACL 规则的报文进行流量统计（**S2700/3700**系列交换机中仅可在入方向上应用**ACL** ），（**S5700/6700**系列交换机既可在入方向应用 **ACL**，又可以在出方向上应用 **ACL** ）。在配置基于 ACL 的报文过滤之前需要配置好相应的ACL规则，但每个调用的**ACL**仅可匹配一个**ACL**规则， 当ACL中包括许多规则时，则必须指出所应用的具体ACL规则编号。

1．在全局或VLAN上应用基于ACL的流量统计

在全局或VLAN上应用基于ACL的流量统计配置也是在系统视图下根据不同S系列交换机选择以下不同命令进行的。

（1）在S2700-52P-EI/2700-52P-PWR-EI/2710SI/3700SI/3700EI系列交换机上

执行 traffic-statistic [vlan vlan-id] inboundacl { { [ipv6] {bas-acl | adv-acl |name acl-name } } | l2-acl | user-acl } [rule rule-id] [by-bytes] 命令对匹配单个ACL规则的出 方向报文进行流量统计。

执行 traffic-statistic [vlan vlan-id] inboundacl { l2-acl |nameacl-name } [rule rule- id] acl { bas-acl | adv-acl |nameacl-name } [rule rule-id] [by-bytes] 命令对同时匹配二层ACL和三层ACL规则的出方向报文进行流量统计。

（2）在除S2700-52P-EI和S2700-52P-PWR-EI系列之外的其他S2700EI系列交换机上

执行 traffic-statistic [vlanvlan-id] inboundacl { { [ipv6] {bas-acl |adv-acl |nameacl-name } } | l2-acl } [rule rule-id] 命令对匹配单个ACL规则的入方向报文进行流量统计。

执行 traffic-statistic [vlan vlan-id] inbound acl {l2-acl |nameacl-name } [rule rule- id] acl { bas-acl | adv-acl

|nameacl-name } [rule rule-id] 命令对同时匹配二层 ACL和三层ACL规则的入方向报文进行流量统计。

（3）在S5700/6700系列交换机上

执行 traffic-statistic [vlan vlan-id] inboundacl { [ipv6] {bas-acl | adv-acl |name acl-name } | l2-acl |user-acl } [rule rule-id] [by-bytes] 命令对匹配单个ACL规则的入方向报文进行流量统计。

执行 traffic-statistic [vlanvlan-id] outboundacl { [ipv6] {bas-acl |adv-acl |name acl- name |l2-acl |user-acl } } [rule rule-id] 命令对匹配单个ACL规则的出方向报文进行流量统计。

执行 traffic-statistic [vlan vlan-id] inbound acl { l2-acl |nameacl-name } [rule rule- id] acl { bas-acl |adv-acl |nameacl-name } [rule rule-id] [by-bytes] 命令同时匹配二层ACL和三层ACL规则的入方向报文进行流量统计。

执行 traffic-statistic [vlan vlan-id] outboundacl { l2-acl |nameacl-name } [rule rule- id] acl { bas-acl |adv-acl |nameacl-name } [rule rule-id] 命令对同时匹配二层ACL和三层ACL规则的出方向报文进行流量统计。

以上traffic-statistic命令中的参数和选项说明如表9-14所示。

表9-14 traffic-statistic命令参数和选项说明

参数	说明
vlan *vlan-id*	指定要应用基于 ACL 的流量统计的 VLAN 编号，取值范围为 1～4 094 的整数
inbound	表示对入方向的报文进行流量统计
outbound	表示对出方向的报文进行流量统计
acl	指定基于 IPv4 ACL 对报文进行流量统计
ipv6	指定基于 IPv6 ACL 对报文进行流量统计
bas-acl	指定基于基本 ACL 对报文进行流量统计的 ACL 编号，取值范围为 2 000～2 999 的整数
adv-acl	指定基于高级 ACL 对报文进行流量统计的 ACL 编号，取值范围为 3 000～3 999 的整数
l2-acl	指定基于二层 ACL 对报文进行流量统计的 ACL 编号，取值范围为 4 000～4 999 的整数
user-acl	指定基于用户自定义 ACL 对报文进行流量统计的 ACL 编号，取值范围为 5 000～5 999 的整数
name *acl-name*	指定基于命名型 ACL 对报文进行过滤。其中，*acl-name* 表示 ACL 的名称，为 1～32 个字符，不支持空格，区分大小写，且要以英文字母 a～z 或 A～Z 开始
rule *rule-id*	指定基于 ACL 中特定规则进行报文流量统计。对于 IPv4 ACL，取值范围为 0～4 294 967 294 的整数；对于 IPv6 ACL，取值范围为 0～2 047 的整数
by-bytes	指定按照字节数量统计。缺省情况下，按照报文数量（packets）进行统计。指定 by-bytes 参数，将按照字节数量进行统计

【示例 1】在VLAN100的入方向，配置基于ACL的流量统计，统计匹配ACL3000中rule1规则的报文数量。

<HUAWEI>system-view

[HUAWEI] traffic-statistic vlan 100 inbound acl 3000 rule 1

2．在接口上配置流量统计

在端口上应用基于 ACL 的流量统计的配置也要根据不同 S 系列交换机选择对应的以下配置命令在具体的接口视图下进行。

（1）在S2700-52P-EI/2700-52P-PWR-EI/2710SI/3700SI/3700EI系列交换机上

执行 traffic-statistic inboundacl { { [ipv6] {bas-acl |adv-acl |nameacl-name } } | l2-acl |user-acl } [rule rule-id] [by-bytes] 命令对匹配单个ACL规则的出方向报文进行流量统计。

执行 traffic-statistic inbound acl { l2-acl |name acl-name } [rule rule-id] acl {bas- acl | adv-acl |nameacl-name } [rule rule-id] [by-bytes] 命令对同时匹配二层 ACL和三层ACL规则的出方向报文进行流量统计。

（2）在除S2700-52P-EI和S2700-52P-PWR-EI系列之外的其他S2700EI系列交换机上

执行 traffic-statistic inbound acl { { [ipv6] {bas-acl | adv-acl |nameacl-name } } | l2-acl } [rule rule-id] 命令对匹配单个ACL规则的入方向报文进行流量统计。

执行 traffic-statistic inboundacl { l2-acl |nameacl-name } [rule rule-id] acl {bas-acl |adv-acl |name acl-name } [rule rule-id] 命令对同时匹配二层ACL和三层ACL规则的入方向报文进行流量统计。

（3）在S5700/6700系列交换机上

执行 traffic-statistic inboundacl { [ipv6] {bas-acl |adv-acl |nameacl-name } | l2-acl |user-acl } [rulerule-id] [by-bytes] 命令对匹配单个ACL规则的入方向报文进行流量统计。

执行 traffic-statistic outbound acl { [ipv6] {bas-acl | adv-acl |nameacl-name } | l2-acl } [rule rule-id] 命令对匹配单个ACL规则的出方向报文进行流量统计。

执行 traffic-statistic inboundacl { l2-acl |nameacl-name } [rule rule-id] acl {bas-acl |adv-acl |name acl-name } [rule rule-id] [by-bytes] 命令对同时匹配二层 ACL 和三层ACL规则的入方向报文进行流量统计。

执行 traffic-statistic outbound acl { l2-acl |nameacl-name } [rule rule-id] acl {bas- acl | adv-acl |nameacl-name } [rule rule-id] 命令对同时匹配二层ACL和三层ACL规则的出方向报文进行流量统计。

【示例 2】在接口GE0/0/1的入方向，配置基于ACL的流量统计功能，统计匹配ACL 3000中 rule 1规则的报文数量。

<HUAWEI>system-view

[HUAWEI] interface gigabitethernet 0/0/1

[HUAWEI-GigabitEthernet0/0/1] traffic-statistic inbound acl 3000 rule 1

3．基于ACL的流量统计管理

配置好基于ACL的流量统计后，可执行以下任意视图display traffic-statistics命令查看设备上基于ACL的报文过滤的流量统计信息。

说明

除 S2700-52P-EI 和 S2700-52P-PWR-EI 系列之外的其他 S2700EI 系列交换机不支持user-acl参数。S2700/3700系列交换机中不支持outbound选项，即不支持出方向的流量统计。

（1）display traffic-statistics [vlan vlan-id | interface interface-type interface-number] { inbound | outbound } [acl { bas-acl | adv-acl | user-acl } [rule rule-id]]

（2）display traffic-statistics [vlan vlan-id | interface interface-type interface-number] { inbound | outbound } [acl {acl-name | l2-acl } [rule rule-id] [acl {bas-acl | adv-acl |acl-name } [rule rule-id]]]

（3）display traffic-statistics interface { inbound | outbound }

（4）display traffic-statistics [vlan vlan-id | interface interface-type interface-number] { inbound | outbound } [acl ipv6 {bas-acl | adv-acl | acl-name } [rule rule-id]]

【示例 3】查看接口GE0/0/1入方向上基于ACL 3009的流量统计信息。输出信息中的字段说明如表9-15所示。

< HUAWEI > system-view

[HUAWEI] display traffic-statistics inbound acl 3009

ACL:3009 Rule:1

matched:0 packets, passed:0 packets, dropped:0 packets

表9-15 display traffic-statistics命令输出信息字段说明

字段	说明
ACL	显示应用的 ACL 的编号
Rule	显示应用的 ACL 规则 ID
matched	显示匹配 ACL 规则的报文数量
passed	显示通过报文的数量
dropped	显示丢弃报文的数量

也可执行以下 reset traffic-statistics用户视图命令清除设备上基于ACL的报文过滤的流量统计信息。

（1）reset traffic-statistics [vlan vlan-id | interface interface-type interface-number] { inbound | outbound } [acl { bas-acl| adv-acl | user-acl } [rule rule-id]]

（2）reset traffic-statistics [vlan vlan-id | interface interface-type interface-number] { inbound | outbound } [acl {acl-name| l2-acl } [rule rule-id] [acl { bas-acl | adv-acl |acl-name } [rule rule-id]]]

（3）reset traffic-statistics { interface | }{ inbound |outbound }

（4）reset traffic-statistics [vlanvlan-id | interface interface-type interface-number] { inbound |outbound } [acl ipv6 {bas-acl | adv-acl | acl-name } [rule rule-id]]

9.4 ACL配置示例

为了使大家对ACL配置和应用有一个全面的理解，下面分别介绍基本ACL、高级ACL、二层ACL和用户自定义ACL的配置方法。

9.4.1 基本ACL配置示例

本示例拓扑结构如图9-1所示，Switch作为FTP服务器（172.16.104.110/24），已知 Switch与各个子网之间路由可达。现要通过基本ACL过滤用户访问FTP服务器时报文中源IP地址，限制用户访问交换机上FTP服务器的权限，具体要求如下。

（1）子网1（172.16.105.0/24）的所有用户在任意时间都可以访问FTP服务器。

（2）子网2（172.16.107.0/24）的所有用户只能在某一个时间范围内访问FTP服务器。

（3）其他用户不可以访问FTP服务器。

说明

在华为S系列交换机的许多种登录（如Telnet、STelnet、HTTP、HTTPS、FTP和 FTPS等）、访问中都可以通过ACL进行用户权限控制，具体参见本书第3章相关内容。

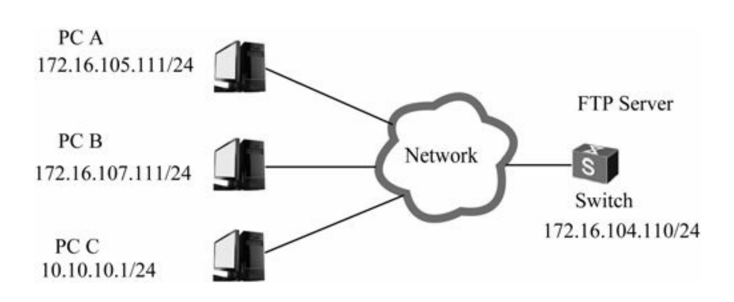


图9-1 基本ACL应用示例拓扑结构

1．基本配置思路分析

本示例的基本配置思路如下。

（1）在Switch上创建基本ACL，并通过三条基本ACL规则中分别对3个子网用户发送的报文中的源IP地

址进行过滤（其实就是一种访问授权）；然后为允许子网2中的用户访问FTP服务器配置一个基本ACL生效时间段。

（2）在Switch上启用FTP功能。

（3）在Switch的FTP服务器上调用上面所创建的基本ACL，对网络中不同用户的FTP服务器访问进行控制。

2．具体配置步骤

下面是具体的配置步骤。

（1）配置用于允许子网2用户访问FTP服务器的ACL生效时间段。假设为从2013年1月1日开始到2013年12月31日，每个双休日的下午2点到下午6点生效。

<HUAWEI > system-view

[HUAWEI] sysname Switch

[Switch] time-range ftp-access from 0:0 2013/1/1 to 23:59 2013/12/31 #---配置时段时间段范围为2013年1月1日0时开始到2013年12月31日23时59分结束

[Switch] time-range ftp-access 14:00 to 18:00 off-day #---配置每个周六、周日的下午 2点到下午 6点的时间段

（2）配置基本ACL。

[Switch] acl number 2001

[Switch-acl-basic-2001] rule permit source 172.16.105.0 0.0.0.255 #---允许子网1用户的报文通过

[Switch-acl-basic-2001] rule permit source 172.16.107.0 0.0.0.255 time-range ftp-access #---允许子网2中的用户报文在名为ftp-access的时间段通过

[Switch-acl-basic-2001] rule deny source any #---禁止其他所有用户的报文通过

[Switch-acl-basic-2001] quit

（3）启用FTP服务器功能。

[Switch] ftp server enable

（4）在FTP服务器的访问中调用前面的基本ACL。

[Switch] ftp acl 2001

9.4.2 高级ACL配置示例

本示例拓扑结构如图9-2所示，Switch为 S5700 系列交换机堆叠系统，是一个高级ACL应用示例。要求禁止研发部门和市场部门在上班时间（8:00至17:30）访问工资查询服务器（IP 地址为 10.164.9.9），而总裁办公室不受限制，可以随时访问。



图9-2 高级ACL配置示例拓扑结构

1．基本配置思路分析

本示例要求控制指定源IP地址的用户访问指定目的IP地址的主机，所以必须配置高级ACL。这里可以采取两种配置方法：一是通过前面介绍的基于ACL的简化流策略，在连接工资查询服务器的交换机 GE2/0/1端口或者两部门各自所连接的GE1/0/2和GE1/0/3端口入方向上应用所配置的高级ACL；二是通过QoS流策略。

当采用基于ACL的简化流策略配置方法时，基本的配置思路如下。

（1）配置ACL生效时间段。

（2）配置所需的两条高级ACL（包括ACL规则）。

（3）在交换机GE1/0/2和GE1/0/3端口入方向上分别应用所配置的对应高级ACL。

如果是采用QoS流策略配置方法，则基本的配置思路如下。

（1）配置ACL生效时间段。

（2）配置所需的两条高级ACL（包括ACL规则）。

（3）配置根据上述高级ACL进行的流分类。

（4）配置对上述流分类采取拒绝的流行为。

（5）配置并在研发部门和市场部门连接的交换机端口上应用QoS流策略。

说明

QoS策略的配置包括定义流分类、定义流行为、创建QoS流策略和应用QoS流策略这四项主要任务，具体将在第10章介绍。

2．具体配置步骤

下面分别介绍以上两种配置方法。

（1）基于ACL的简化策略的配置方法

① 配置ACL生效时间段。

[HUAWEI] time-range satime 8:00 to 17:30 working-day　#---配置在工作日的 8:00至 17:30的时间段

② 配置两条ACL及其规则。

[HUAWEI] acl 3002

[HUAWEI-acl-adv-3002] rule deny ip source 10.164.2.0 0.0.0.255 destination 10.164.9.9 0.0.0.0 time-range satime　#---配置禁止市场部门访问工资查询服务器的访问规则

[HUAWEI-acl-adv-3002] quit

[HUAWEI] acl 3003

[HUAWEI-acl-adv-3003] rule deny ip source 10.164.3.0 0.0.0.255 destination 10.164.9.9 0.0.0.0 time-range satime #---配置禁止研发部门到工资查询服务器的访问规则

[HUAWEI-acl-adv-3003] quit

③ 在端口上应用ACL。

如果采用第一种方法，需分别在交换机GE1/0/2和GE1/0/3端口入方向上应用所配置的高级ACL。可随便选择使用traffic-filter或者traffic-secure命令对三层报文进行过滤。下面仅以traffic-filter命令为例进行介绍。

[HUAWEI] interface GigabitEthernet 1/0/2

[HUAWEI-GigabitEthernet1/0/2] traffic-filter inbound acl 3002

[HUAWEI-GigabitEthernet1/0/2] quit

[HUAWEI] interface GigabitEthernet 1/0/3

[HUAWEI-GigabitEthernet1/0/3] traffic-filter inbound acl3003

[HUAWEI-GigabitEthernet1/0/3] quit

（2）基于QoS流策略的配置方法

如果采用第二种方法，首先也需要按照前面介绍的第①和第②步配置好 ACL 生效时间段和两条高级 ACL，然后还需要按照以下的步骤配置好流分类、流行为和流策略，最后在两部门连接的交换机端口入方向上应用对应的流策略。

① 定义基于ACL的流分类。

#---配置流分类 c_market，对匹配ACL 3002的报文进行分类

[HUAWEI] traffic classifier c_market

[HUAWEI-classifier-c_market] if-match acl 3002

[HUAWEI-classifier-c_market] quit

#---配置流分类 c_rd，对匹配ACL 3003的报文进行分类。

[HUAWEI] traffic classifier c_rd

[HUAWEI-classifier-c_rd] if-match acl 3003

[HUAWEI-classifier-c_rd] quit

② 定义流行为。

#---定义流行为b_market，动作为拒绝报文通过。

[HUAWEI] traffic behaviorb_market

[HUAWEI-behavior-b_market] deny

[HUAWEI-behavior-b_market] quit

#---定义流行为b_rd，动作为拒绝报文通过。

[HUAWEI] traffic behaviorb_rd

[HUAWEI-behavior-b_rd] deny

[HUAWEI-behavior-b_rd] quit

③ 创建流策略，对以上流分类和流行为进行关联。

#---配置流策略p_market，将流分类c_market与流行为b_market关联。

[HUAWEI] traffic policy p_market

[HUAWEI-trafficpolicy-p_market] classifier c_marketbehaviorb_market

[HUAWEI-trafficpolicy-p_market] quit

#---配置流策略p_rd，将流分类c_rd与流行为b_rd关联。

[HUAWEI] traffic policy p_rd

[HUAWEI-trafficpolicy-p_rd] classifier c_rd behaviorb_rd

[HUAWEI-trafficpolicy-p_rd] quit

④ 在对应端口上应用流策略。

#---将流策略p_market应用到GE1/0/2接口。

[HUAWEI] interface gigabitethernet 1/0/2

[HUAWEI-GigabitEthernet1/0/2] traffic-policy p_market inbound

[HUAWEI-GigabitEthernet1/0/2] quit

#---将流策略p_rd应用到GE1/0/3接口。

[HUAWEI] interface gigabitethernet 1/0/3

[HUAWEI-GigabitEthernet1/0/3] traffic-policy p_rd inbound

[HUAWEI-GigabitEthernet1/0/3] quit

以上就是两种方法的全部配置。配置完成后可以用display acl all命令查看ACL配置信息，以验证配置结果。

9.4.3 二层ACL配置示例

本示例如图9-3所示，Switch作为网关设备，下挂用户PC。现要求配置ACL，禁止源MAC地址为00e0-f201-0101、目的MAC地址为0260-e207-0002的报文通过。

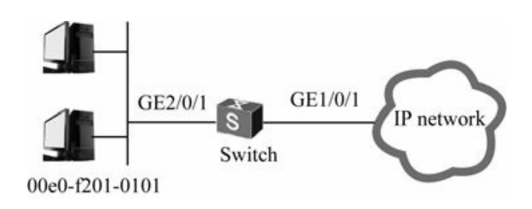


图9-3 二层ACL配置示例拓扑结构

1．基本配置思路分析

本示例同样有两种配置方法，即基于 ACL的简化策略法和 QoS 流策略法。下面分别予以介绍其基本配置思路。

如果采用基于ACL的简化策略配置方法，则基本配置思路如下。

（1）配置所需的二层ACL（包括ACL规则）。

（2）在交换机GE2/0/1接口上应用上面所配置的二层ACL。

如果采用的是QoS流策略配置方法，则基本配置思路如下。

（1）配置所需的二层ACL（包括ACL规则）。

（2）然后同样是QoS流策略的四大配置任务：定义基于以上二层ACL的流分类，定义一个所需的流行为，创建一个QoS流策略，将前面定义的流分类和流行为关联起来，在交换机接口上应用所创建的QoS流策略。

2．具体配置步骤

下面同样介绍以上两种配置方法的具体配置步骤。

（1）基于ACL的简化策略的配置方法

① 配置符合要求的二层ACL。因为要匹配的仅是一个MAC地址，所以源MAC地址和MAC地址的掩码

均为0xffff-ffff-ffff。

<HUAWEI>system-view

[HUAWEI] acl 4000

[HUAWEI-acl-L2-4000] rule deny source-mac 00e0-f201-0101 ffff-ffff-ffffdestination-mac 0260-e207-0002 ffff-ffff-ffff

[HUAWEI-acl-L2-4000] quit

②在GE2/0/1端口上应用以上配置的二层ACL 4000。同样，可随便选择使用 traffic-filter或者traffic-secure命令对二层报文进行过滤。下面仅以traffic-secure命令为例进行介绍。

[HUAWEI] interface GigabitEthernet 2/0/1

[HUAWEI-GigabitEthernet2/0/1] traffic-secure inbound acl 4000

[HUAWEI-GigabitEthernet2/0/1] quit

（2）基于QoS流策略的配置方法

首先也是要按照上面介绍的配置好二层ACL及其规则。然后进行以下配置。

① 配置流分类 tc1，对匹配ACL 4000的报文进行分类。

[HUAWEI] traffic classifier tc1

[HUAWEI-classifier-tc1] if-match acl 4000

[HUAWEI-classifier-tc1] quit

② 定义流行为tb1，动作为拒绝报文通过。

[HUAWEI] traffic behavior tb1

[HUAWEI-behavior-tb1] deny

[HUAWEI-behavior-tb1] quit

③ 创建流策略tp1，将流分类tc1与流行为tb1关联。

[HUAWEI] traffic policy tp1

[HUAWEI-trafficpolicy-tp1] classifier tc1behavior tb1

[HUAWEI-trafficpolicy-tp1] quit

④ 将流策略tp1应用到GE2/0/1接口。

[HUAWEI] interface gigabitethernet 2/0/1

[HUAWEI-GigabitEthernet2/0/1] traffic-policy tp1 inbound

[HUAWEI-GigabitEthernet2/0/1] quit

以上就是本示例的全部配置步骤。

9.4.4 用户自定义ACL配置示例

本示例拓扑结构如图9-4所示，Switch的GE1/0/1接口连接用户，GE2/0/1接口连接上层路由器。要求在接口GE1/0/1下绑定用户自定义ACL，从二层报文头偏移14个字节开始匹配，拒绝匹配成功的报文通过，匹配 的字符串内容为0x0180C200（共4个字节）。

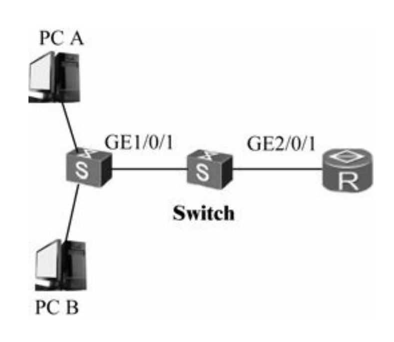


图9-4 用户自定义ACL配置示例拓扑结构

下面同样以基于 ACL 的简化策略和基于QoS 策略这两种配置方法为例进行介绍，配置思路与前面两节介绍的配置示例中的配置思路一样，不同的只是所配置的ACL类型不同，本示例所配置的是用户自定义ACL。下面直接介绍具体的配置步骤。

1．基于ACL的简化策略的配置步骤

（1）配置符合要求的用户自定义ACL。

```
[HUAWEI] acl 5000
[HUAWEI-acl-user-5000] rule deny l2-head 0x0180C200 0xffffffff 14
[HUAWEI-acl-user-5000] quit
```

（2）在 Switch 的 GE1/0/1 端口上应用上面创建的用户自定义 ACL。下面仅以traffic-secure命令为例进行介绍。

```
[HUAWEI] interface GigabitEthernet 1/0/1
[HUAWEI-GigabitEthernet1/0/1] traffic-secure inbound acl 5000
[HUAWEI-GigabitEthernet1/0/1] quit
```

2．基于QoS流策略的配置步骤

首先也是按照上面介绍的配置好二层ACL及其规则。然后进行以下配置。

（1）配置流分类 tc1，对匹配ACL 5000的报文进行分类。

```
[HUAWEI] traffic classifier tc1
[HUAWEI-classifier-tc1] if-match acl 5000
[HUAWEI-classifier-tc1] quit
```

（2）定义流行为tb1，动作为拒绝报文通过。

```
[HUAWEI] traffic behavior tb1
[HUAWEI-behavior-tb1] deny
[HUAWEI-behavior-tb1] quit
```

（3）创建流策略，将流分类与流行为关联。

```
[HUAWEI] traffic policy tp1
[HUAWEI-trafficpolicy-tp1] classifier tc1behavior tb1
[HUAWEI-trafficpolicy-tp1] quit
```

（4）在接口GE1/0/1下应用流策略。

```
[HUAWEI] interfacegigabitethernet 1/0/1
[HUAWEI-GigabitEthernet1/0/1] traffic-policy tp1 inbound
[HUAWEI-GigabitEthernet1/0/1] quit
```

以上就是本示例的全部配置步骤。

9.5 自反ACL

自反ACL（Reflective ACL）是动态ACL技术的一种应用。它根据 IP报文的上层会话信息生成，只有当私网用户先访问了公网后才允许对应的公网用户访问本地私网。利用自反ACL可以很好地保护企业内部网络免受外部非法用户的攻击。但要注意，只能针对高级**ACL**或者高级**ACL6**进行自反**ACL**，并且只能根据**TCP**、**UDP**和**ICMP**协议类型 的报文自动生成ACL规则，在华为S系列交换机中仅**S7700/9300/9700**系列支持 。

9.5.1 自反ACL的基本工作原理

自反ACL有以下两个显著的特点。

（1）由内网始发的流量到达配置了自反ACL功能的设备后，设备根据此流量的第三层和第四层信息自动生成一个临时性的反向ACL，并保持一段时间。此临时性ACL规则中的协议类型不变，源IP地址和目的IP地址，以及源端口与目的端口与始发ACL规则对调。

（2）当对端设备发出的响应报文到达配置了自反 ACL 功能的设备时，会自动根据这个临时性的ACL允许响应通信通过。那么设备是如何确定该响应通信是始发ACL通信的响应通信呢？它是依据响应报文中的第三、四层信息与先前始发ACL通信报文中的第三、四层信息是否严格匹配来判定的。不完全匹配的不允许访问，这样既保证了外网响应流量通过，又拒绝了非法的外网用户主动访问。

根据不同的匹配规则，自反ACL的实现原理如下。

（1）当配置自反ACL功能的接口通过TCP或UDP协议类型的报文时，接口将下发一条报文源IP地址和目的IP地址、源端口和目的端口互换的ACL规则。

（2）当配置自反ACL功能的接口通过ICMP协议类型的报文时，自反ACL功能阻止目的端发送的ICMP-Echo-Request报文通过，允许接收目的端发送的ICMP-Echo-Reply报文。

（3）自反ACL匹配的协议类型与触发接口自动生成自反ACL的报文协议类型相同。

如图9-5所示，在交换机上配置自反ACL功能后，外网无法主动访问内网。这时，一个源 IP IPa，源端口Porta，目的 IP IPb，目的端口Portb的报文发往外网，设备会自动生成一条自反ACL的规则，允许源 IP IPb，源端口Portb，目的 IP IPa，目的端口Porta的报文通过。

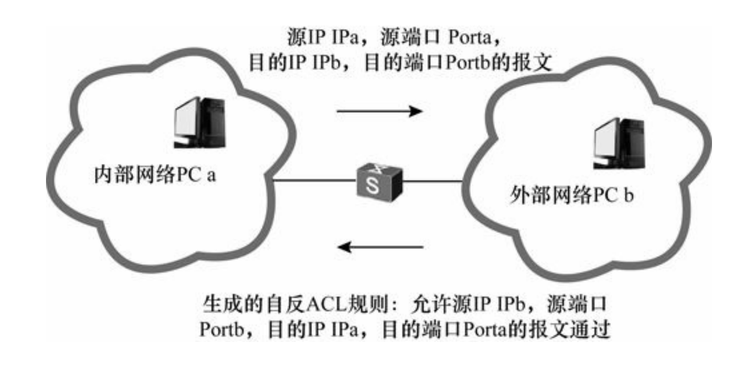


图9-5 自反ACL工作原理示例

9.5.2 配置自反ACL

自反 ACL 可以很好地保护企业内部网络，免受外部非法用户的攻击。在整个自反ACL的配置中可以配置的任务包括以下3个方面。

（1）配置需要用于启用自反ACL功能的高级ACL。

（2）在对应交换机端口上启用对应高级ACL的自反ACL功能，并可选择配置该自反ACL的老化时间。

（3）（可选）在交换机上全局配置自反ACL的老化时间。

以上3项配置任务的具体配置步骤如表9-16所示。

表9-16 自反ACL的配置步骤

<table>
<tr><th>配置任务</th><th>步骤</th><th>命令</th><th colspan="2">说明</th></tr>
<tr><td rowspan="5">配置高级ACL</td><td>1</td><td>system-view
例如：<HUAWEI> system-view</td><td colspan="2">进入系统视图</td></tr>
<tr><td>2</td><td>time-range time-name { start-time to end-time days | from time1 date1 [to time2 date2] }
例如：[HUAWEI] time-range test 14:00 to 18:00 off-day</td><td colspan="2">（可选）创建一个 ACL 生效的时间段。命令中的参数说明参见 9.2.1 节表 9-4 中的第 2 步</td></tr>
<tr><td rowspan="2">3</td><td>acl [number]acl-number [match-order { auto | config }]
例如：[HUAWEI] acl number 3100</td><td>使用编号创建一个数字型的高级 ACL 并进入高级 ACL 视图。命令中的参数和选项说明参见 9.2.2 节</td><td rowspan="2">二选一</td></tr>
<tr><td>acl name acl-name { basic | acl-number } [match-order { auto | config }]
例如：[HUAWEI] acl name test1 3100</td><td>使用名称创建一个命名型的高级 ACL 并进入高级 ACL 视图。命令中的参数和选项说明参见 9.3.2 节</td></tr>
<tr><td>4</td><td>当参数 protocol 为 TCP 时使用以下命令创建高级 ACL 规则：
rule [rule-id] { deny | permit } { protocol-number | tcp } [destination { destination-address destination-wildcard | any } |destination-port { eq port | gt port | lt port | range port-start port-end } | { { precedence precedence | tos tos } * | dscpdscp } | fragment | logging | source { source-address source-wildcard | any } | source-port { eq port | gt port | lt port |range port-start port-end } | tcp-flag { ack | fin | psh | rst | syn | urg } * | time-range time-name | ttl-expired] *
例如：[HUAWEI-acl-adv-3100] rule permit tcp destination 10.1.1.0 0.255.255.255 destination-port eq 80 source 192.168.1.0 0.0.0.255 source-port eq 8080 time-range test</td><td colspan="2">（三选一）配置高级 ACL 规则，命令中的参数和选项说明参见 9.2.2 节
参数 dscp dscp 和 precedence precedence 不能同时配置；参数 dscp dscp 和 tos tos 不能同时配置</td></tr>
</table>

（续表）

<table>
<tr><th>配置任务</th><th>步骤</th><th>命令</th><th>说明</th></tr>
<tr><td>配置高级ACL</td><td>4</td><td>当参数 protocol 为 UDP 时使用以下命令创建高级 ACL 规则：
rule [rule-id] { deny | permit } { protocol-number | udp } [destination { destination-address destination-wildcard | any } |destination-port { eq port | gt port | lt port | range port-start port-end } | { { precedence precedence | tos tos } * | dscpdscp } | fragment | logging | source { source-address source-wildcard | any } | source-port { eq port | gt port | lt port |range port-start port-end } | time-range time-name | ttl-expired] *
例如：[HUAWEI-acl-adv-3100] rule permit udp destination 10.1.1.0 0.255.255.255 destination-port eq 42 source 192.168.1.0 0.0.0.255 source-port eq 42 time-range test

当参数 protocol 为 ICMP 时使用以下命令创建高级 ACL 规则：
rule [rule-id] { deny | permit } { protocol-number | icmp } [destination { destination-address destination-wildcard | any } | { { precedence precedence | tos tos } * | dscp dscp } | fragment | logging | icmp-type { icmp-name | icmp-type icmp-code } |source { source-address source-wildcard | any } | time-range time-name | ttl-expired] *
例如：[HUAWEI-acl-adv-3100]rule permit icmp destination 10.1.1.0 0.255.255.255 source 192.168.1.0 0.0.0.255 time-range test</td><td>（三选一）配置高级 ACL 规则，命令中的参数和选项说明参见 9.2.2 节
参数 dscp dscp 和 precedence precedence 不能同时配置；参数 dscp dscp 和 tos tos 不能同时配置</td></tr>
<tr><td rowspan="3">配置自反 ACL 功能</td><td>5</td><td>quit
例如：[HUAWEI-acl-adv-3100] quit</td><td>退出高级 ACL 视图，返回系统视图</td></tr>
<tr><td>6</td><td>interface interface-type interface-number
[HUAWEI] interface gigabitethernet 1/0/0</td><td>进入需要配置自反 ACL 功能的接口视图。自反 ACL 需要在接口上进行配置，接口对报文进行过滤</td></tr>
<tr><td>7</td><td>traffic-reflect { inbound |outbound }acl { adv-acl-name | adv-acl-number } [timeout time-value]
例如：[HUAWEI-GigabitEthernet1/0/0] traffic-reflect outbound acl 3000</td><td>使能自反 ACL 功能和配置自反 ACL 老化时间。命令中的参数和选项说明如下。
（1）inbound：二选一选项，指定该接口为内网接口。如果只希望对外网用户访问某个内网用户的权限进行限制，则需要在连接该内网用户的接口上配置自反 ACL 并选择此选项
（2）outbound：二选一选项，指定该接口为外网接口。如果一个外网接口对应多个内网用户，希望对外网用户访问所有内网用户的权限进行限制，则需要在该连接外网的接口上配置自反 ACL，并选择此选项
（3）adv-acl-name：二选一参数，指定一个要启用自反 ACL 功能的高级 ACL 名称。为 1～32 个字符，不支持空格，区分大小写，且要以英文字母 a～z 或 A～Z 开始；可以是英文字母、数字和"#"、"%"、"-"等字符的组合
（4）adv-acl-number：二选一参数，指定一个要启用自反 ACL 功能的高级 ACL 编号，取值范围为 3 000～3 999
（5）time-value：可选参数，指定自反 ACL 的老化时间，取值范围为（60～2 147 483）整数秒。使能自反 ACL 功能之后，缺省情况下，接口下的自反 ACL 老化周期是下面可选配置的全局自反 ACL 老化周期
缺省情况下，自反 ACL 功能未使能，可用 undo traffic-reflect { inbound | outbound } acl { adv- acl-name | adv-acl- number }命令去使能自反 ACL 功能</td></tr>
<tr><td rowspan="2">（可选）配置全局自反 ACL 老化时间</td><td>8</td><td>quit
例如：[HUAWEI-GigabitEthernet1/0/0]quit</td><td>退出接口视图，返回系统视图</td></tr>
<tr><td>9</td><td>traffic-reflect timeout time-value
例如：[HUAWEI] traffic-reflect timeout 6000</td><td>配置全局自反 ACL 老化周期，取值范围为（60～2 147 483）整数秒
【说明】如果已经使用 traffic-reflect 命令在接口视图下配置了自反 ACL 老化周期，则以接口视图下配置的老化周期为准；如果在接口视图下没有配置自反 ACL 老化周期，则以 traffic- reflect timeout 命令在系统视图下配置的老化周期为准
如果在老化周期内有符合自反 ACL 规则的报文通过接口，该接口的自反 ACL 规则被保留。如果在老化周期内没有符合自反 ACL 规则的报文通过接口，该接口的自反 ACL 规则被删除
当报文流量较大时，可以适当减小自反 ACL 的老化周期，增大老化的频率；当报文流量较小时，可以适当增大自反 ACL 的老化周期，减小老化的频率
缺省情况下是没有配置全局自反 ACL 老化周期的，可用 undo traffic-reflect timeout 命令取消原来的全局自反 ACLf 老化周期配置</td></tr>
</table>

配置好后可以使用display traffic-reflect { inbound |outbound } [interface interface- type interface-number] [acl {adv-acl-name |adv-acl-number }] 命令查看配置自反ACL的信息。

【示例 1】在GE1/0/0端口出方向上配置自反ACL功能，对所有TCP报文进行自反。

```
<HUAWEI>system-view
[HUAWEI] acl 3000
[HUAWEI-acl-adv-3000] rule permit tcp
[HUAWEI-acl-adv-3000] quit
[HUAWEI] interface gigabitethernet 1/0/0
[HUAWEI-GigabitEthernet1/0/0] traffic-reflect outbound acl 3000
```

【示例 2】查看ACL 3001在出方向上的自反ACL信息。

```
<HUAWEI>display traffic-reflect outbound acl 3001
Proto   SP    DP   DIP        SIP        Count   Timeout   Interface
-----------------------------------------------------------------------------
UDP     2     80   1.1.1.1    2.2.2.2    9       5(s)      Eth-Trunk2
-----------------------------------------------------------------------------
* Total <1> flows accord with condition, <1> items was displayed,
-----------------------------------------------------------------------------
* Proto=Protocol,SIP=Source IP,DIP=Destination IP,Timeout=Time to cutoff,
* SP=Source port,DP=Destination port,Count=Packets count(data).
```

9.5.3 自反ACL配置示例

本示例拓扑结构如图 9-6 所示， Switch 的 GE1/0/1 端口连接了内网的用户， GE2/0/1 端口连接到 Internet。在GE2/0/1 端口的出方向上配置基于 UDP协议的自反ACL功能，当内网的主机先访问 Internet 中的服务器之后才允许Internet的服务器访问内网的主机。同时，在全局和GE2/0/1端口下配置自反ACL的老化时间，对自反ACL进行自动老化。

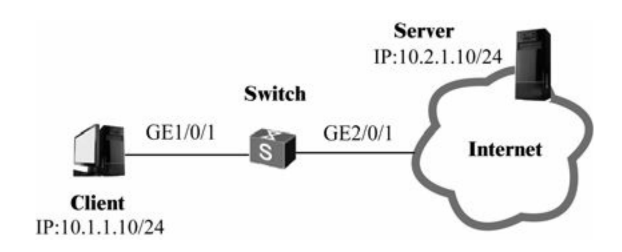


图9-6 自反ACL配置示例拓扑结构

根据上节介绍的配置任务和配置步骤，可很容易得出本示例的配置步骤，具体如下。

（1）配置高级ACL，允许UDP报文通过。

```
<HUAWEI>system-view
[HUAWEI] acl 3000
[HUAWEI-acl-adv-3000] rule permit udp
[HUAWEI-acl-adv-3000] quit
```

（2）在 GE2/0/1 端口出方向上自反 ACL 功能和老化时间（假设为 600s），对 UDP报文进行自反。

[HUAWEI] interface gigabitethernet 2/0/1

[HUAWEI-GigabitEthernet2/0/1] traffic-reflect outbound acl 3000 timeout 600

[HUAWEI-GigabitEthernet2/0/1] quit

（3）配置全局自反ACL老化时间。但此时在GE2/0/1端口出方向上的自反ACL的老化时间仍是在该端口上配置的老化时间——600s。

[HUAWEI] traffic-reflect timeout 900

配置完成后，可以通过display traffic-reflect任意视图命令查看自反ACL信息、验证配置结果。具体示例如下。从输出信息可以看出，在GE2/0/1端口下对UDP协议的报文进行了自反，并且对自反后的报文进行统计，自反ACL的老化时间为600s。

[HUAWEI] display traffic-reflect outbound acl 3000

Proto SP DP DIP SIP Count Timeout Interface

UDP 2 80 10.2.1.10 10.1.1.10 9 600(s) GigabitEthernet2/0/1

----

* Total <1> flows accord with condition, <1> items was displayed,

----

* Proto=Protocol,SIP=Source IP,DIP=Destination IP,Timeout=Time to cutoff,

* SP=Source port,DP=Destination port,Count=Packets count(data).

第10章　QoS基础及技术原理

QoS（质量服务）是一项非常复杂的技术，但它的应用又非常广泛。它可实现的主要功能包括流量监管（对进入接口的，超出限制速率的报文进行丢弃）、流量整形（对接口发送的，超出限制速率的报文先进行缓存，等待流量不超出速率时发送）、拥塞避免（在出现网络拥塞时对符合条件的报文进行丢弃）、拥塞管理（在出现网络拥塞时采用队列调度的方法对符合条件的队列中的报文优先发送）、流策略（可根据不同的流分类实现诸如禁止/允许通过，重标记报文优先级、重标记报文VLAN标签、重定向流量、过滤报文、流量镜像、启用流量统计功能等行为）。

本章将详细介绍以上各种QoS基础知识和技术原理，为下一章介绍的各种QoS功能和QoS流策略应用配置方法打下坚实的基础。

10.1 QoS基础

QoS（Quality of Service，服务质量）是一种可以为不同类型业务流提供差分（即“不同”）服务等级的技术。通过 QoS 可以给那些对带宽、时延、时延抖动、丢包率等敏感的业务流提供更加优先的服务等级，使这些业务能满足用户正常、高性能使用的需求。

10.1.1 QoS概述

在传统的IP网络中，所有的报文都被无区别地同等对待。即每个网络设备对所有的报文均采用FIFO（First In First Out，先入先出）的策略进行处理，依照报文到达时间的先后次序分配所需要的资源，尽最大的努力（Best-Effort）将报文送到目的地。但在这种方式下，对报文传送的可靠性、传送延迟、丢包率等性能都不提供任何保证，所以仅适用于对这些服务性能不敏感的普通业务，如WWW、FTP文件传输、E-mail等业务。

随着IP互联网上新型应用的不断出现，对IP网络的服务质量也提出了新的要求，比如远程教学、远程医疗、可视电话、电视会议、视频点播等。在这些对实时性和连续性方面要求更加苛刻的应用中，如果报文传送延时太长将是用户无法接受的，因为在这类应用中是不能容忍中间停顿的现象的。为了支持具有不同服务需求的话音、视频以及数据等业务，要求网络能够区分出不同的业务类型，进而为之提供相应等级的服务。QoS正是这样一种可以为不同业务类型报文提供差分服务的技术，通过对网络流量进行调控，可避免并管理网络拥塞，减少报文丢包率。

QoS服务等级就是指对业务流所需的带宽、时延、时延抖动、丢包率等核心需求的评估。当然，不同类型的业务所需要评估的因素并不一样，如普通数据流在带宽、丢包率方面要求更高，而像视频通信之类的业务流则在时延和时延抖动方面要求更高。

1．带宽

“带宽”又可称为吞吐量，表示在一定时间内业务流的平均速率，单位通常是kbit/s。QoS可以为不同业

务流分配不同的端口带宽，以实现高优先级，或者对带宽需求更高的业务流（如视频流等）分配到更大的端口带宽，实现更加快速的数据传输。

2．时延

“时延”表示业务流穿过网络时需要的平均时间。对于网络中的一个设备来说，一般将时延的需求理解为几种等级。通过优先队列（端口有几个优先级不同的数据发送队列）的调度方法使得高优先级的队列业务尽可能快地获得传输服务，而低优先级的队列业务则需要等待没有高优先级业务时才能获得传输服务。

3．时延抖动

“时延抖动”表示业务流穿过网络的时间的变化。不同的业务流对时延抖动的敏感度也不一样，像话音、视频这类实时要求比较高的业务要求时延抖动更小，否则就可能出现话音、视频流断断续续，不连续或者失真的现象。

4．丢包率

“丢包率”表示业务流在传送过程中的丢失比率。由于现代的传输系统具有很高的可靠性，信息的丢失往往发生在网络出现拥塞时。同样不同业务流对丢包率的敏感度也 不一样，此时如话音、视频流对丢包率就不是很敏感了，其中丢掉几帧视频是看不出太大的变化的，而对于数据文件来说就非常敏感了，因为这样可能导致数据最终无法使用。

说明

在QoS的分类流程中最关键的是对各种不同业务流配置不同的优先级，对流入设 备的业务流按其优先级进行分类，然后为不同类型业务流定义一个相应的流行为，设备就会为对应的业务流执行相应的QoS行为。

不同的报文使用不同的QoS优先级，例如二层VLAN报文使用802.1p优先级，三层IP报文使用DSCP优先级，MPLS报文使用EXP优先级（本章不介绍此优先级）。下面两小节将分别介绍二层的802.1p优先级，三层的IP优先级和DSCP优先级。

10.1.2 二层VLAN帧中的优先级

二层帧中的优先级是专门针对VLAN帧的，因为普通二层帧中是不携带优先级字段的。VLAN帧中的优先级那就是我们通常所说的 802.1p优先级（由 IEEE 802.1p协议定义），位于VLAN帧中的“802.1Q Tag”字段的“PRI”子字段中，如图 10-1所示。

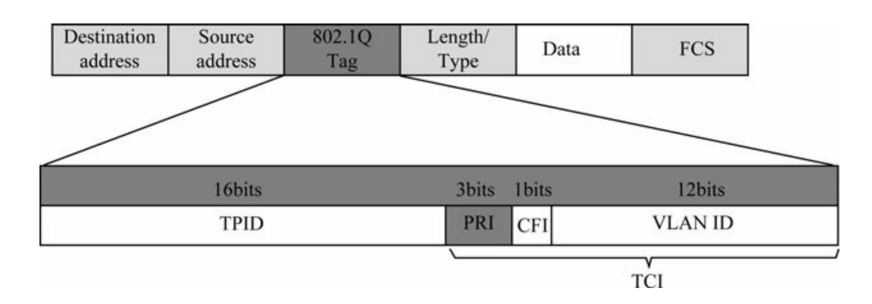


图10-1 VLAN帧中的802.1p优先级字段

IEEE 802.1p是 IEEE 802.1Q（VLAN标签技术）标准的扩充协议，它们协同工作。IEEE 802.1p的出现，使得第二层交换机能够提供流量优先级和动态组播过滤服务，其中流量优先级规范工作在媒体访问控制（MAC）层，组播流量过滤功能可确保该流量不超出第二层交换网络范围。

IEEE 802.1Q标准定义了为以太网 MAC 帧添加的标签，但并没有定义和使用优先级字段，而使用IEEE

802.1p修改后的以太网MAC帧的以太网协议头中则定义了该字段。802.1p优先级位于二层VLAN帧头部，适用于不需要分析三层报文头，而需要在二层环境下保证QoS的场合。4个字节的 802.1Q标签头包含了 2个字节的TPID（Tag Protocol Identifier，标签协议标识，取值为 0x8100）和 2个字节的TCI（Tag Control Information，标签控制信息），参见图10-1。

TCI部分中PRI子字段就是802.1p优先级，也称为CoS优先级。它由3位组成，取值范围为0～7，共可表示8个优先级。其中，最高优先级为7，应用于网络管理和关键性网络流量，如路由选择信息协议（RIP）和开放最短路径优先（OSPF）协议的路由表更新；优先级6和5主要用于延迟敏感（delay-sensitive）应用程序，分别对应交互式话音和视频；优先级 4 到 1 主要用于受控负载（controlled-load）应用程序、流式多媒体（streaming multimedia）、关键性业务流量（business-critical traffic），如SAP数据和后 台流量。优先级0是缺省值，并在没有设置其他优先级值的情况下自动启用。

10.1.3 三层IP报文中的优先级

上面介绍的二层VLAN帧优先级比较简单，就是由PRI子字段的三位来标识，共有8种优先级，但在三层IP报文中，优先级的描述就要复杂许多，并且在不同时期出现了两种不同的优先级类型和不同的标识方法。

1．ToS字段标的IP优先级

在早期的RFC 791标准中，IP数据包是依赖ToS（Type of Service，服务类型）字段来标识数据优先级值的。ToS是IP数据包中的IP报头中的一个字段（共1个字节），用来指定IP包的优先级，设备会优先转发ToS值高的数据包。

ToS字段共一个字节（8位），包括3个部分：0～2共3位用来定义数据包的IP优先级（IP Precedence）、ToS和最后一个固定为0的位，如图10-2所示。

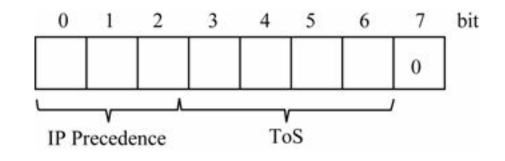


图10-2 IP包头中的ToS字段结构

（1）IP Precedence部分

IP优先级部分共3位，取值范围为0～7（值越大，优先级越高）。用名称表示时，这8个取值分别 为routine（普通，值为000）、priority（优先，值为001）、immediate（快速，值为010）、flash（闪速，值为011）、flash-override（急速，值为100）、critical（关键，值为101）、internetwork control（网间控制，值为 110）和network control（网络控制，值为 111），分别对应于数字0～7。

在以上IP优先级值中，6和7一般保留给网络控制数据使用，比如路由；5推荐给话音数据使用；4推荐由视频会议和视频流使用；3推荐给话音控制数据使用；1和2推荐给数据业务使用；0为缺省标记值。在IP优先级配置时，既可以使用0～7这样的数值，也可以使用上述对应的优先级名称。

（2）ToS

在IP包头的ToS字段中紧接着IP优先级字段后面的四位是ToS部分，代表需要为对应报文提供的服务类型（标识报文所注重的特性要求）。一开始，在RFC 791中只用到了第3～5位，分别代表IP包在Delay（延时），Throughput（吞吐量），Reliability（可靠性）这三方面的特性要求（每个报文在这三位中只有一位可能置1，此时表示IP包在对应方面有特别要求）。后来在RFC1349标准中又扩展到第6位，表示IP包在路径

开销（cost）方面的特性要求。

要注意的是，虽然ToS部分共有4位，但每个IP包中这4位中只能有一位为1，所以实际只有5个取值（包括全为0的值）。这5个值所对应的名称和数值分别为normal （一般服务，取值为 0000）、min-monetary-cost（最小开销，取值为 0001，确保路径开销最小）、max-reliability（最高可靠性，0010，确保可靠性最高）、max-throughput（最大吞吐量，取值为0100，确保传输速率最高）、min-delay（最小时延，取值为1000，确保传输时延最小）。

2．DS字段的DSCP优先级和PHB

在后来新的RFC 2474标准中，重新定义了原来 IP包头部的ToS字段，并改称为DS（Differentiated Services，差分服务）字段，也是共一个字节（8位）。总的来说，第0～5位 （共 6位）用来表示DSCP（Differentiated Services Code Point，差分服务代码点）优先级，取值范围为0～63，共能标识出 64 个优先级值（值越大，优先级越高），最后两位保留，用于显示拥塞通知（Explicit Congestion Notification，ECN），如图10-3所示。

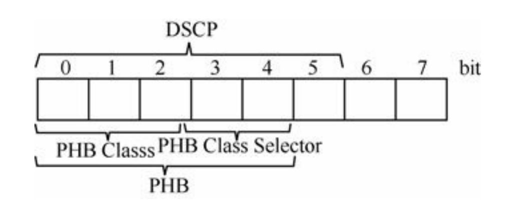


图10-3 IP包头中的DS字段结构

后来在 IETF RFC 2597标准中定义了 PHB （Per-Hop Behavior，逐跳行为），通过PHB值可以确定在网关处对 IP包的转发行为。这个PHB值是通过前面介绍DSCP优先级部分的第0～4位来标识的，其中第0～2位用来标识 PHB类别（PHB Class）值，共 8个值，对应表示为CS0～CS7，对应在RFC 791定义的 8个IP优先级值，而第 3～4位用来标识PHB类别选择（PHB Class Selector）值，参见图10-3。PHB类别值和PHB类别选择值共同组成PHB值。DSCP值是由PHB的5位再加上第5位（固定为0），但在PHB类别中的3位不能全为**0** 。

在RFC 2597中定义了4种确保转发（Assured Forwarding，AF）PHB组（称为AF PHB）。它使用了DS字段中的第0～2位定义PHB类别，而使用DS字段中的第3和4位代表报文的“丢弃优先级”，用AF（x，y）表示，其中x表示流分类，y表示对应的丢弃优先级。

说明

所谓“确保转发”就是允许管理员在没有超过线路允许速率的情况下提供尽可能 的传输质量保证，但如果超出用户线路速率则可能在出现拥塞时丢弃数据包。

在确保转发PHB中，定义了4种PHB类别（也即“流分类”），它们的值分别为001、010、011和100（对应CS1～CS4），它们本身代表了流的不同优先级（值越大转发优先级越高 ），然后通过第3和4位的丢弃优先级值（取非0的3个值，分别为01、10和11，值越大丢弃优先级越高 ）进一步区分同一类流不同IP包的丢弃优先级。它们共同针对4种PHB分类组成了4组AF等级，它们所对应的AF值和对应的DSCP值如表10-1所示（此时第5位的值固定为0）。

表10-1 4组AF PHB等级

丢弃优先级	Class 1	Class 2	Class 3	Class 4
低丢弃优先级	AF11 (DSCP 10)：001010	AF21 (DSCP 18)：010010	AF31 (DSCP 26)：011010	AF41 (DSCP 34)：100010
中丢弃优先级	AF12 (DSCP 12)：001100	AF22 (DSCP 20)：010100	AF32 (DSCP 28)：011100	AF42 (DSCP 36)：100100
高丢弃优先级	AF13 (DSCP 14)：001110	AF23 (DSCP 22)：010110	AF33 (DSCP 30)：011110	AF43 (DSCP 38)：100110

再后来在RFC 3246标准中，又定义一个加速转发（Expedited Forwarding，EF）PHB，对应CS5，即在DS字段中的第0～2位取值为101，第3～4位取值固定为11，第5位固定为0，这样一来对应的DSCP值就为46（101110）。EF PHB具有低延时、低开销和低抖动特性，适用于话音、视频和其他实时服务，一般具有比其他通信类型更加优先的队列。

除了前面介绍的AF和EF外，还有一个缺省的PHB，那就是尽力服务类型，它所对应的DSCP值为000000，即十进制的0。另外还定义了CS6和CS7，CS6用于网间控制，对应的DSCP为110000，即十进制的48；CS7用于网内控制，对应的DSCP值为 111000，即十进制的56。

在配置DSCP优先级时，既可以使用对应的DSCP名称，如CS6、CS7、AF11、AF12 （在CS1～CS4中每个包含了一组DSCP值，所以要指定具体的DSCP名称），又可使用对应的DSCP十进制值，如48、56等。

3．IP优先级与DSCP优先级的对应关系

DSCP优先级是向后兼容IP优先级的，当支持DSCP的设备收到仅支持ToS中的IP优先级的报文时，缺省情况下它们之间是有一种映射关系的，具体如表10-2所示。当然，如果设备仅支持ToS的IP优先级，缺省情况下是不能识别报文中的DSCP优先级值的，这时需要事先在接收设备配置好DSCP优先级与IP优先级的映射关系。这方面具体在本章后面介绍。

表10-2 IP优先级与DSCP优先级值的对应关系

3位IP优先级的值	对应的ToS字段高6位						对应的6位DSCP优先级	3位IP优先级的值	对应的ToS字段高6位						对应的6位DSCP优先级
	7	6	5	4	3	2			7	6	5	4	3	2	
0	0	0	0	0	0	0	0	4	1	0	0	0	0	0	32
	0	0	0	0	0	1	1		1	0	0	0	0	1	33
	0	0	0	0	1	0	2		1	0	0	0	1	0	34
	0	0	0	0	1	1	3		1	0	0	0	1	1	35
	0	0	0	1	0	0	4		1	0	0	1	0	0	36
	0	0	0	1	0	1	5		1	0	0	1	0	1	37
	0	0	0	1	1	0	6		1	0	0	1	1	0	38
	0	0	0	1	1	1	7		1	0	0	1	1	1	39
1	0	0	1	0	0	0	8	5	1	0	1	0	0	0	40
	0	0	1	0	0	1	9		1	0	1	0	0	1	41
	0	0	1	0	1	0	10		1	0	1	0	1	0	42
	0	0	1	0	1	1	11		1	0	1	0	1	1	43
	0	0	1	1	0	0	12		1	0	1	1	0	0	44
	0	0	1	1	0	1	13		1	0	1	1	0	1	45
	0	0	1	1	1	0	14		1	0	1	1	1	0	46
	0	0	1	1	1	1	15		1	0	1	1	1	1	47
2	0	1	0	0	0	0	16	6	1	1	0	0	0	0	48
	0	1	0	0	0	1	17		1	1	0	0	0	1	49
	0	1	0	0	1	0	18		1	1	0	0	1	0	50
	0	1	0	0	1	1	19		1	1	0	0	1	1	51
	0	1	0	1	0	0	20		1	1	0	1	0	0	52
	0	1	0	1	0	1	21		1	1	0	1	0	1	53
	0	1	0	1	1	0	22		1	1	0	1	1	0	54
	0	1	0	1	1	1	23		1	1	0	1	1	1	55
3	0	1	1	0	0	0	24	7	1	1	1	0	0	0	56
	0	1	1	0	0	1	25		1	1	1	0	0	1	57
	0	1	1	0	1	0	26		1	1	1	0	1	0	58
	0	1	1	0	1	1	27		1	1	1	0	1	1	59
	0	1	1	1	0	0	28		1	1	1	1	0	0	60
	0	1	1	1	0	1	29		1	1	1	1	0	1	61
	0	1	1	1	1	0	30		1	1	1	1	1	0	62
	0	1	1	1	1	1	31		1	1	1	1	1	1	63

10.1.4 三种QoS服务模型

"服务模型"就是设备为不同业务流提供服务的一种模式。总体来说，包括 Best Effort、IntServ和DiffServ三种服务模型。下面分别予以简单介绍。

1．Best Effort模型

Best Effort（尽力而为）模型是一种为所有业务流提供相同服务等级的服务模型，也是最简单的服务模型。在Best Effort模型中，应用程序可以在任何时候发出任意数量的报文，而且不需要事先获得批准，也不需要通知网络，网络尽最大的可能性发送每一个数据报文，但对时延、可靠性等性能不提供任何保证。

Best Effort模型是 Internet的缺省服务模型，它适用于绝大多数网络，如FTP、E-mail等，它通过先进先出（FIFO）调度方式来实现。

2．IntServ模型

IntServ（Integrated Service，综合服务）模型的主要特点是在发送报文前要先向网络提出申请。这个请求是通过协议信令来完成的，如RSVP（Resource Reservation Protocol，资源预留协议）。应用程序首先通过RSVP信令通知网络它的QoS需求（如时延、带宽、丢包率等指标），在收到资源预留请求后，传送路径上的网络节点实施许可控制（Admission control），验证用户的合法性并检查资源的可用性，决定是否为应用程序预留资源。一旦认可并为应用程序的报文分配了资源，则只要应用程序的报文控制在流量参数描述的范围内，网络节点将承诺满足应用程序的 QoS 需求。传输路径上的网络节点可以通过执行报文的分类、流量监管、低延迟的排队调度等行为，来满足对应用程序的承诺。

IntServ模型常与组播应用结合，适用于需要保证带宽、低延迟的实时多媒体应用，如电视会议、视频点播等。当前，采用RSVP协议的IntServ 模型定义了两种业务类型。

（1）保证型服务（Guaranteed Service）：提供保证的带宽和时延限制来满足应用程序的要求。如VoIP（Voice over IP，IP话音）应用可以预留10MB带宽和要求不超过1s的时延。

（2）负载控制型服务（Controlled-Load Service）：保证即使在网络过载（overload）的情况下，仍能对报文提供类似Best Effort模型在未过载时的服务质量，保证某些应用程序报文的低时延和低丢包率需求。

IntServ模型的最大优点是可以提供端到端的QoS传输服务，最大缺点是可扩展性不好：网络节点需要为每个资源预留维护一些必要的软状态（Soft State）信息；在与组播应用相结合时，还要定期地向网络发资源请求和路径刷新信息，以支持组播成员的动态加入和退出。而这些操作要耗费网络节点较多的处理时间和内存资源。在网络规模扩大时，维护的开销会大幅度增加，对网络节点特别是核心节点线速处理报文的性能造成严重影响。因此，IntServ模型不适宜于在流量汇集的骨干网上大量应用。

3．DiffServ 模型

为了在 Internet 上针对不同的业务提供有差别的服务， IETF 定义了 DiffServ （Differentiated Service，差分服务）模型。

DiffServ模型是一种多服务模型，可以满足不同用户业务流的QoS需求。它与IntServ模型不同的是应用程序在发出报文前通过设置报文头部的优先级字段，向网络中各设备通告自己的QoS需求，而不需要通知途经的网络设备为其预留资源，网络不需要为每个流维护状态，仅根据每个报文携带的优先级就可确定所需为对应流提供的服务等级。

DiffServ模型一般用来为一些重要的应用提供端到端的QoS，这也是本章介绍的QoS技术中所使用的服务模式。通常在配置DiffServ模型后，边界设备通过报文的源地址和目的地址等信息对报文进行分类，对不同的报文设置不同的优先级，并标记在报文头部， 而其他设备只需要根据设置的优先级来进行报文的调度。

在DiffServ模型中，流分类、流量监管、流量整形、拥塞管理和拥塞避免是对不同类型报文提供有区别服务的基石，它们主要完成如下功能。

（1）流分类：依据一定的匹配规则识别出不同类型的报文，是有区别地实施服务的前提。

（2）流量监管：对进入交换机的特定流量的规格进行监管。当流量超出规格时，可以采取限制或惩罚措施，以保护运营商的商业利益和网络资源不受损害。

（3）流量整形：一种主动调整流的输出速率的措施，使流量适配下游交换机可供给的网络资源，避免不必要的报文丢弃和拥塞。

（4）拥塞管理：在发生网络拥塞时必须采取解决资源竞争的措施。通常是将报文放入队列中缓存，并采取某种调度算法安排报文的转发次序。

（5）拥塞避免：过度的拥塞会对网络资源造成损害。拥塞避免功能可以监督网络资源的使用情况，在发现拥塞有加剧的趋势时采取主动丢弃报文的策略，通过调整流量措施来解除网络的过载。

在这些功能组件中：流分类是基础，它依据一定的匹配规则识别出报文，是有区别地实施服务的前提；流量监管、流量整形、拥塞管理和拥塞避免从不同方面对网络流量及其分配的资源实施控制，是有区别地提供服务具体体现。

10.1.5 DiffServ模型体系结构

DiffServ体系结构定义了实现差分服务的系统模型和基本功能组件，如图10-4所示。在一个网络节点上，实现差分服务的基本功能组件包括“逐跳行为”（PHB）、业务流分类和流量调整（包括流量监管与流量整形、拥塞避免与拥塞管理）等功能。差分服务建立在DS（差分服务）域模型之上，并规定了一个DS域的边界节点和内部节点。在DS域边界节点上，对进入网络的业务流进行分类、流量调整和优先级标记，并按照 DS 域所支持的PHB组中的一个PHB进行转发。在内部节点上，将根据边界节点标记的DSCP或802.1p优先级所定义的PHB来选择该业务流的转发行为，为业务流分配带宽资源。

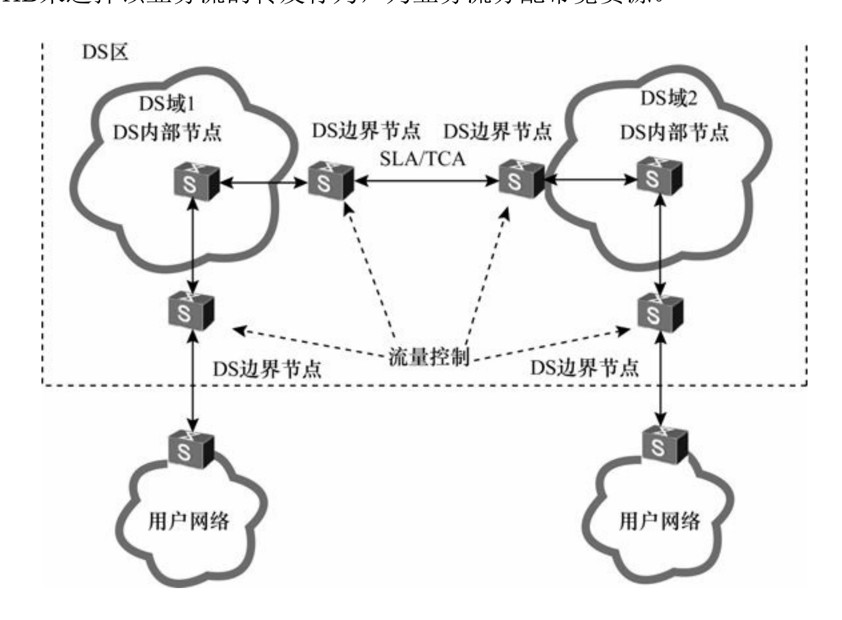


图10-4 DiffServ模型体系结构

1．DS域

DiffServ模型的实现基于DS域，DS域由一组采用相同的服务提供策略和实现了相同PHB组集合的相连DS节点组成。一个DS域由DS边界节点和DS内部节点组成，边界节点构成了DS域的边界，内部节点构成了DS域的核心。

2．DS节点

DS节点指实现DiffServ功能的网络节点。DS节点可分为DS边界节点和DS内部节点。

（1）DS边界节点

DS边界节点负责连接另一个DS域或者连接一个没有DS功能的域的节点。DS边界节点负责将进入此 DS 域的业务流进行分类和可能的流量调整，以保证穿过此 DS 域的业务流被适当标记，并按照DS域所支持的PHB组中的一个PHB进行转发。

对于不同方向的业务流，DS边界节点既可以是DS域的输入（Ingress）节点，又可以是DS域的输出（Egress）节点。业务流在Ingress节点处进入DS域，在Egress节点处离开DS域。Ingress节点负责保证进入DS域的业务流符合本域和此节点直连的另一个域之间的SLA（Service Level Agreements，服务等级协定）或TCA（Traffic Conditioning Agreement，流量控制协定）。Egress节点依据两个域之间的TCA细节，对转发到其直连的对等域的业务流执行流量调整功能。

（2）DS内部节点

DS内部节点负责连接同一DS域中的其他DS内部节点或DS边界节点。DS内部节点负责根据IP报头中的DS字段或VLAN报文的802.1p字段所定义的PHB来为该业务流选择转发行为。无论是DS边界节点还是DS内部节点都必须根据业务流的DSCP或者802.1p选择相应的PHB进行转发操作。

3．SLA

SLA（服务等级协定）指用户（个人、企业、有业务往来的相邻ISP等）和服务提供商签署的关于业务流在网络中传递时所应当获得的待遇。SLA 包括很多方面，例如付费协议，其中的技术说明部分称为SLS（Service Level Specification，服务等级规范）。

4．TCA

TCA（流量控制协定）指用户与服务提供商签署的关于业务分类准则、业务模型及相应处理的协定。去掉了商业条款的TCA称为TCS（Traffic Conditioning Specification，流量控制规范），一个SLA中可以包含TCA。对于业务的处理而言，SLA或SLS指明的是比较一般的内容，例如采用什么样的机制，而TCA或TCS则比较具体，例如具体的带宽要求。

5．DS区

一个或多个邻接的DS域统称为DS区。DS区可以支持贯穿区内多个DS域的分类业务。DS区中的DS域可能支持不同的PHB组，和QoS优先级到PHB的映射规则。不同DS域可有不同的PHB，以实现不同的服务提供策略。它们之间通过SLA和TCA协调提供跨区域服务。SLA/TCA指明了如何在DS域边界节点调整从一个DS域传向另一个DS域的业务流。

10.2 QoS优先级映射

优先级映射用来实现报文携带的 QoS 优先级与设备内部优先级（又称为本地优先级，是设备为报文分配的具有本地意义的优先级）之间的转换，从而使设备根据内部优先级为不同报文提供有差别的QoS服务质量。不同S系列交换机所支持的优先级信任模式和优先级映射模式都有所不同，下面将具体介绍。

10.2.1 优先级映射

不同网络中的报文使用不同类型的 QoS 优先级字段，例如 VLAN 网络中的报文使用802.1p优先级，IP网络中的报文使用IP优先级或DSCP优先级。当报文经过不同网络时，为了保持报文的优先级，需要在连接不同网络的设备上配置这些优先级字段的映射关系。当设备连接不同网络时，所有进入设备的报文优先级

（包括802.1p和DSCP，统称为“外部优先级”）字段根据所配置优先级映射表都被转换为交换机端口的内部优先级；当设备发出报文时，又需要根据优先级映射表将报文中携带的内部优先级恢复为原来的对应外部优先级。

从以上分析可以看出，这里涉及到设备信任报文的哪种优先级，也就是下面即将介绍的优先级信任模式，以及对应的内、外部优先级映射模式。

1．优先级信任模式

配置优先级信任模式可以确定设备根据哪种优先级进行映射。相当于你说你有多种身份，但最终要确定以哪种身份来衡量你的地位。

（1）三种优先级信任模式

在华为S系列交换机中，有以下三种优先级信任模式。

① 信任报文的 802.1p优先级：配置信任 802.1p优先级时，设备根据报文的 802.1p优先级对报文进行分类并进行后续的优先级映射，得到报文映射后的802.1p优先级，IP优先级，或DSCP优先级。缺省情况下，端口信任报文的802.1p优先级。

② 信任报文的 DSCP优先级：配置信任 DSCP优先级时，设备根据报文的 DSCP优先级对报文进行分类并进行后续的优先级映射，得到报文映射后的802.1p优先级，或DSCP优先级，或丢弃优先级（只有支持**DSCP**优先级的**IP**报文才可映射丢弃优先级，因为只有这类报文才有这样的数据位，参见10.1.3节介绍）。

③ 信任报文的 IP优先级：配置交换机端口信任 IP优先级，此时将按照报文中所携带的IP优先级查找对应的IP优先级映射表，得到报文映射后的802.1p优先级或IP优先级。

（2）华为S系列交换机对优先级信任模式的支持

在不同的华为S系列交换机中，所支持的优先级信任模式不完全一样，具体表现如下。

①S2700SI/S2700EI/S2710SI系列交换机仅支持“信任报文的802.1p优先级”模式。

② S5700HI/5710EI/6700/7700/9300/9700系列交换机仅支持“信任报文的802.1p优先级”和“信任报文的DSCP优先级”这两种模式。

③ S2700-52P-EI/2700-52P-PWR-EI/3700SI/3700EI/S5700SI/5700EI/5700LI/5700S-LI 系 列交换机全面支持以上3种优先级信任模式。

说明

在报文进入设备端口之后，如果报文携带了VLAN标签，则可以选择信任802.1p优先级；如果报文没有携带VLAN标签，则报文会根据端口缺省的802.1p优先级进行转发，该端口优先级即报文转发时进入的端口队列号。在三层转发时，可以选择信任DSCP优先级。

2．优先级映射模式

为了保证不同报文的服务质量，对于进入设备的报文，设备可以根据配置将报文携带的优先级映射为内部优先级，并根据内部优先级与队列之间的映射关系确定报文进入的队列，从而针对队列进行流量整形、拥塞避免、队列调度等处理；报文从设备发送出去时，设备可以根据配置修改报文发送出去时所携带的优先级，以便其他设备根据报文的优先级提供相应的QoS服务。配置优先级映射模式可以确定报文优先级与内部优先级的映射关系，以便设备在后续转发中根据内部优先级提供有差别的QoS服务。

（1）优先级映射模式

总体来说，华为S系列交换机所支持的所有优先级映射模式包括以下几种（不考虑MPLS报文中的EXP优先级）。

① 内部优先级和队列之间的映射

② DSCP优先级到802.1p、新的DSCP优先级、丢弃优先级的映射。

③ IP优先级到802.1p、新的IP优先级的映射。

④ 802.1p优先级到PHB行为/颜色的映射。

⑤ PHB行为/颜色到802.1p优先级的映射。

⑥ DSCP优先级到PHB行为/颜色的映射。

⑦ PHB行为/颜色到DSCP优先级的映射。

（2）华为S系列交换机对优先级信任模式的支持

同样，不同的华为S系列交换机所支持的优先级映射模式也不完全相同。

① 在S2700SI/2700EI/2710SI系列交换机中仅支持“内部优先级和队列之间的映射”模式。

② 在S2700-52P-EI/2700-52P-PWR-EI/S3700SI/3700EI/5700SI/5700EI/5700LI/5700S-LI系列交换机中支持以下3种优先级映射模式。

● 内部优先级和队列之间的映射。

● DSCP优先级到 802.1p、新的DSCP优先级、丢弃优先级的映射。

● IP优先级到 802.1p、新的 IP优先级的映射。

③ 在 S5700HI/5710EI/6700/7700/9300/9300E/9700系列交换机中，支持以下几种优先级映射模式。

● 在接口入方向，将802.1p优先级映射为PHB行为/颜色。

● 在接口出方向，将PHB行为/颜色映射为802.1p优先级。

● 在接口入方向，将DSCP优先级映射为PHB行为/颜色。

● 在接口出方向，将PHB行为/颜色映射为DSCP优先级。

说明

在 S5700HI/5710EI/6700/7700/9300/9300E//9700系列交换机中支持 RFC 2597和 RFC 3246标准中的PHB（参见本章 10.1.3节），用户可以根据DiffServ域中定义的报文优先级与 PHB 行为/颜色之间的映射关系对报文进行分类。对于来自上游设备的报文，在设备的入接口上绑定DiffServ域，在DiffServ域中将报文携带的优先级信息映射到相应的PHB行为/颜色，然后根据内部优先级与队列之间的映射关系确定报文进入的队列；在设备的出接口上，根据报文的PHB行为进行拥塞管理，根据报文的颜色进行拥塞避免；然后对于流向下游设备的报文，在设备的出接口上绑定DiffServ域，在DiffServ域中将报文的PHB行为/颜色映射为相应的优先级，下游设备根据报文的优先级提供相应的QoS服务。

10.2.2 内部优先级与802.1p和入队列索引的映射关系

“内部优先级”是指设备对报文进行处理的优先级，报文中携带的优先级只有在与内部优先级进行了映射后，才能体现报文在对应设备的最终优先级，设备对报文的处理优先级不是直接按照报文中所携带的各种优先级进行的，而是按照设备中配置的内部优先级进行的。根据10.1.3节介绍的不同PHB行为定义了8种“内部优先级”，从低到高的顺序依次是：BE、AF1、AF2、AF3、AF4、EF、CS6 和 CS7（不区分大小写）。

1．802.1p优先级与内部优先级的映射关系

缺省情况下，所有华为S系列交换机的802.1p优先级与内部优先级的映射关系是一样的，如表 10-3 所示。从中可以看出，这些交换机中 802.1p 优先级与内部优先级的缺省映射关系是按等级一一对应的，即最低的802.1p优先级0对应最低的内部优先级BE，最高的802.1p优先级7对应最高的内部优先级CS7。

表10-3 缺省的802.1p优先级与内部优先级映射关系

802.1p 优先级	内部优先级
0	BE
1	AF1
2	AF2
3	AF3
4	AF4
5	EF
6	CS6
7	CS7

2．内部优先级与入队列索引的映射关系

在实际部署时有时需要调整服务等级与入端口队列的映射关系，或者将不同的服务等级放入同一入端口队列中进行调度，从而有效地节约设备缓存。设备按照内部优先级将报文送入不同的入端口队列，从而针对队列进行流量整形、拥塞避免、队列调度等处理。但不同S系列交换机的内部优先级与入队列的索引关系也有所不同。

说明

内部优先级与入队列的映射配置仅在 S2700SI/3700/5700/6700 系列交换机上支 持，其他系列均仅可采用缺省映射关系。

不同 S 系列交换机中内部优先级与入队列索引之间的映射关系有所不同。在 S2700-52P-EI/2700-52P-PWR-EI/3700SI/3700EI/5700HI/5710EI 系列交换机中，缺省情况下，8个内部优先级与8个入队列索引号之间是由低到高（指优先级和入队列索引号）一一对应的映射关系，如表10-4左边部分所示。但在S2700SI系列，以及除S2700-52P-EI和S2700-52P-PWR-EI之外的其他S2700EI系列交换机中，因为仅支持4个入队列，所以每两个内部优先级映射同一个入队列（如BE和AF1这两个最低的内部优先级同时映射最小的0号队列，而CS6和CS7这两个最高的内部优先级同时映射最大的3号队列），但总体上也是由低到高一一对应的，如表10-4右边部分所示。

表10-4 S2700/3700/5700系列交换机缺省的内部优先级与入队列索引映射关系

S2700-52P-EI/2700-52P-PWR-EI/3700SI/3700EI/5700HI/5710EI 系列缺省的内部优先级与入队列索引映射关系		S2700SI，以及除 S2700-52P-EI 和 S2700-52P-PWR-EI 之外的其他 S2700EI 系列缺省的内部优先级与入队列索引映射关系	
内部优先级	入队列索引	内部优先级	入队列索引
BE	0	BE	0
AF1	1	AF1	0
AF2	2	AF2	1
AF3	3	AF3	1
AF4	4	AF4	2
EF	5	EF	2
CS6	6	CS6	3
CS7	7	CS7	3

在S6700系列，以及S7700系列的ES1D2X40SFC0单板、ES1D2L02QFC0单板，S9300系列的LE0DX40SFC00单板、LE1D2L02QFC0单板和S9300E的LE0DX40SFC00单板和LH2D2L02QFC0单板，S9700系列的EH1D2X40SFC0单板、EH1D2L02QFC0单板中内部优先级与各队列之间的对应关系如表 10-5 所示。从中可以看出，内部优先级和入队列的映射关系还要看报文是否是已知的单播报文，如果是已知单播报文，则与前面介绍的表10-4左边部分完成一样，只是针对非已知单播报文情况下的映射关系有所不同。

表10-5 S6700系列，以及S7700/9300/9700系列部分单板缺省的内部优先级与入队列索引映射关系

内部优先级	队列索引	内部优先级	队列索引
BE（非已知单播报文）	0	BE（已知单播报文）	0
AF1（非已知单播报文）	1	AF1（已知单播报文）	1
AF2（非已知单播报文）	1	AF2（已知单播报文）	2
AF3（非已知单播报文）	1	AF3（已知单播报文）	3
AF4（非已知单播报文）	2	AF4（已知单播报文）	4
EF（非已知单播报文）	2	EF（已知单播报文）	5
CS6（非已知单播报文）	6	CS6（已知单播报文）	6
CS7（非已知单播报文）	6	CS7（已知单播报文）	7

在 S5700SI/5700EI/5700LI/5700S-LI/5700HI/5710EI 子系列，以及 S7700/9300/9700系列交换机其他单板中内部优先级与各队列之间的对应关系与表10-4左边部分相同，也是由低到高（指优先级和入队列索引号）一一对应的。

10.3 流量监管和流量整形

如果不限制用户发送的业务流量，大量用户不断突发的业务数据会使网络更加拥挤。为了使有限的网络资源能够更好地发挥效用，更好地为更多的用户服务，必须对用户的业务流量加以限制。本节将要介绍的流量监管（TP，Traffic Policing）和流量整形（TS， Traffic Shaping）就可以通过监督进入网络的流量速率来限制流量及其资源的使用。进行流量监管和流量整形有一个前提条件，就是要判断流量是否超出了规格，然后才能根据评估结果实施调控策略。在流量监管和流量整形功能实现中，一般采用令牌桶（Token Bucket）对流量的规格进行评估。下面先具体了解一下“令牌桶”技术原理。

10.3.1 QoS令牌桶基本工作原理

这里的“令牌桶”是指网络设备的内部存储池（也就是用于缓存数据的内存），而“令牌”（Token）则是指以给定速率填充令牌桶的虚拟信息包。“令牌桶”可以简单理解为一个水桶，而“令牌”则可以理解为通过一根水管流到水桶中的水。

交换机在入端口接收每个帧时都将一个令牌添加到令牌桶中，但这个令牌桶底部有一个孔，不断地按你指定作为平均通信速率（单位为bit/s）的速度领出令牌（也就是从桶中删除令牌的意思），其实就是不断地从出端口发送数据的过程。相当于一个水桶的上边连接一根进水的水管，而下边又连接一根到用水的地方的出水管。在每次向令牌桶中添加新的令牌包时，交换机都会检查令牌桶中是否有足够容量（也就是在要向桶中加水前，先要检查桶内是否已满了），如果没有足够的空间，包将被标记为不符规定的包，这时在包上将发生指定监管器中规定的行为（丢弃或标记）。就相当于如果当前水桶满了，但上边水管的水还是来了，这时要么让这些水白白流到桶外，要么把这些水用其他容器先装起来，等水桶中不再水满时再倒进去，供用户使用。

最初的令牌桶模型考虑的是单令牌桶结构，这个令牌桶称为CBS（Committed Burst Size，承诺突发尺寸），简称C桶，而向C桶中填充令牌的平均速率称之为CIR（Committed Information Rate，承诺信息速率），如果用T_c 表示当前令牌桶中的令牌数，则这个单令牌桶的基本工作原理可以用图10-5来表示。用文字描述如下（假设用B来表示新接收的数据包大小）：

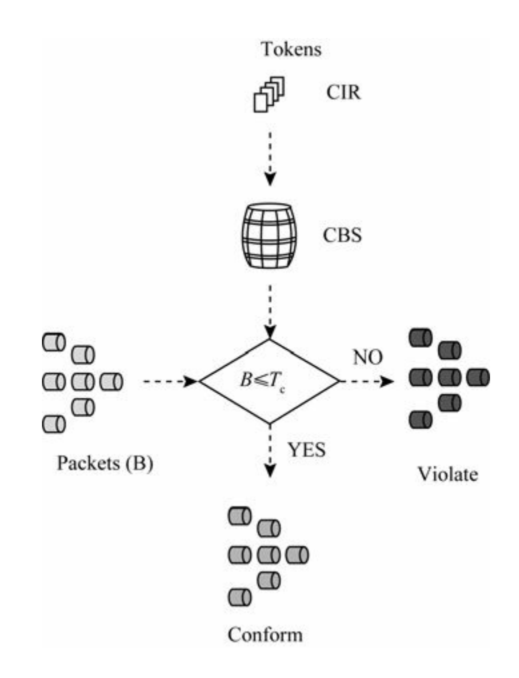


图10-5 令牌桶的基本工作作原理

（1）系统按照CIR速率向令牌桶（相当于交换体机缓存）中投放令牌（相当于从端口上接收到数据包）。

（2）当T_c <CBS时再比较B与T_c 的大小关系。如果B≤T_c，则表示新接收的数据包大小符合规定（Conform）可完整地被缓存到令牌桶中，此时数据包将被标记为绿色（表示允许转发的），当前的T_c 值要相应减少B。

（3）如果B>T_c，则表示新接收的数据包大小违规（Violate）不能完整地被缓存到令牌桶中，此时整个数据包被标记为红色，并将被直接丢弃，T_c 值不减少。

令牌桶填满的时间长短是由令牌桶深度（也就是交换机的缓存大小，单位为 bit，类似于水桶的深度）、令牌漏出速率（类似桶下边接的水管的水速）和超过平均速率的突发流量（类似于上桶上边水管突发的急速水流）持续时间这3个方面共同决定。令牌桶的大小是通过“突发时长上限”乘以“点对点传输时的帧数限制”得出（也就类似突发水流持续的时间*突发水流的流速）。如果突发时间比较短，令牌桶不会溢出，在通信流上不会发生行为；但是如果突发时间比较长，并且速率比较高，令牌桶将溢出，这时将对突发过程中的帧采取相应的流监管策略行为（也就是在水桶水满后对溢出的水的处理方法）。

10.3.2 单速率三色标记算法

在令牌桶处理包的行为方面，主要包括两种令牌桶算法：RFC 2697定义的单速率三色标记（srTCM，single rate three color marker）算法和RFC 2698定义的双速率三色标记（trTCM，two rate three color marker）算法，其评估结果都是为包打上红、黄、绿三色标记（所以称为“三色标记”，有关这些颜色的具体含义将在具体算法中介绍）。QoS会根据包的颜色，设置包的丢弃优先级，其中单速率三色标记比较关心包尺寸的突发，而双速率三色标记则关注速率上的突发，两种算法都可工作于色盲模式和非色盲模式（具体在下面介绍）。本节先介绍单速率三色标记算法原理。

srTCM（单速率三色标记）是一种“单速双桶”算法，它可对流量进行测评，根据评估结果为报文打颜色标记，即绿色、黄色和红色。这里首先要理解“单速”是指算法中的两个令牌桶有同样的承诺信息速率

（CIR），也就是具有相同平均访问速率；这两个令牌桶分别是正常使用的令牌桶（也就是下面将要说到的C桶）和超出令牌桶容量的突发令牌桶（也就是下面将要说到的 E 桶），可以理解为两个水桶，一个是正常使用的水桶，另一个是当正常使用的水桶满后用于装多余水的水桶，如图10-6所示。下面具体解释单速率三色标记算法原理。

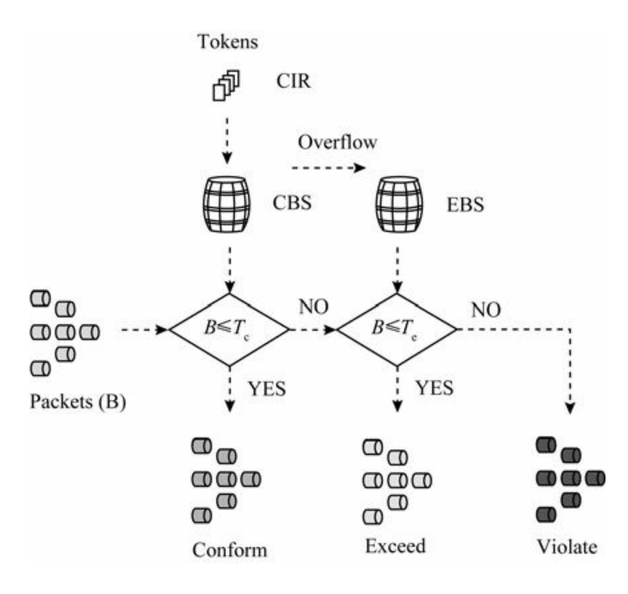


图10-6 单速双桶示意图

1．srTCM算法的3个参数

srTCM算法关注的是数据包的突发尺寸，数据包的色标记评估依据以下3个参数。

（1）承诺信息速率（CIR，Committed Information Rate)：表示向C桶中填充令牌的平均速率，即C桶允许传输或转发报文的平均速率。

（2）承诺突发尺寸（CBS，Committed Burst Size）：表示C桶的容量，即指每次突发所允许的最大的流量尺寸，也相当于允许的最大取令牌的速率，等于桶的容量（最大时一个包就可以全部领取桶中的全部令牌）。

（3）超额突发尺寸（EBS，Excess Burst Size）：表示E桶的容量，即每次突发允许超出CBS的最大流量尺寸。

单速率三色机制采用双桶结构：C桶和E桶（之所以用这两个字母来表示，为的就是与前面说的CBS和EBS两种速率的头个字母一致，便于描述），且两个令牌桶的CIR一样。当C令牌桶满时，超出的令牌也会放在E令牌桶中。

T_c 和T_e 分别表示C令牌桶和E令牌桶中的令牌数，也就是桶中当前的容量（单位也为bit），两桶的总容量分别为CBS和EBS，也就是对应前面介绍的承诺突发尺寸和超额突发尺寸，最初它们都是满的，即T_c 和Te初始值分别等于CBS和EBS。正常情况下不会使用第二个令牌桶（也就是E桶），只有当C令牌桶满后，后面来的令牌才放到E令牌桶中，为出现的突发数据提供信用令牌（也就是经过允许的令牌）。

2．srTCM算法原理

在 srTCM 算法中，两个令牌桶中令牌的添加是按照相同的 CIR 速率进行的。即每隔1/CIR时间添加一个令牌。添加的顺序是先添加C桶再添加E桶，当两个令牌桶中的令牌都满时，再产生的令牌就会被丢弃。系统按照CIR速率向桶中填充令牌。

（1）若T_c <CBS，则T_c 增加。

（2）若T_c =CBS，T_e <EBS，则T_e 增加。

（3）若T_c =CBS，T_e =EBS，则都不增加。

对于到达的报文，用B表示报文的大小。

（1）若$B \le T_c$，报文被标记为绿色，且T_c 减少B。

（2）若$T_c < B \le T_e$，报文被标记为黄色，且T_e 减少B。

（3）若T_e <B，报文被标记为红色，且T_c 和T_e 都不减少。

3．srTCM算法中的报文着色处理

在发送数据包时，令牌的使用IEEE又定义了3种颜色（分别为红色、黄色和绿色）以及两种模式：色盲（color-blind）模式和感色（color-aware）模式，缺省为色盲模式。3种颜色的功能与我们日常生活中的交通指示灯中的3种颜色类似，红色表示违规数据，直接丢弃，黄色表示数据包虽然违法，但不直接丢弃，而是延迟发送，绿色为合法数据包，直接发送。

在色盲（color-blind）模式下假设包都是没有经过“着色”处理的（不辨别包中原来标记的颜色），是根据包长度来确定包被标记的颜色。现假设到达的包长度为B（单位为bit）。若包长度B小于C桶中的令牌数T_c（也就是C桶中的令牌数足够该包发送所需），则包被标记为绿色，表示包符合要求，包发送后C桶中的令牌数T_c 减少B。如果$T_c < B < T_e$（也就是包长度大于C桶中的令牌数，而小于E桶中的令牌数），则标记为黄色，则从E桶中取出所需令牌，E桶中的令牌数T_e 减少B；若$B > T_e$，标记为红色，表示是违反规定 的包，直接丢弃，两令牌桶中的总令牌数都不减少。

在感色（color-aware）模式下是假设包在此之前已经过“着色”处理（会辨别包中原来标记的颜色），如果包已被标记为绿色，或包长度 $B < T_c$（注意只要满足其中一个条件即可，下同），则包被标记为绿色，C桶中的令牌数 T_c 值随之也相应减少 B；如果包已被标记为黄色，或$T_c < B < T_e$，则包被标记为黄色，同时E桶中的令牌数T_e 也随之相应减少B；如果包已被标记为红色，或$B > T_e$，则包被标记为红色，T_c 和T_e 都不减少。

10.3.3 双速率三色标记算法

trTCM（双速率三色标记）是一种双桶双速算法，也可对流量进行测评，根据评估结果为报文打颜色标记，即绿色、黄色和红色。这里同样首先要搞清楚“双速率”是什么意思，它是指该算法中两个令牌桶中的CIR速率不同，存在两个令牌填充速率。

与单速率三色标记算法不同，双速率三色标记算法中的两个令牌桶是C桶和P桶（不是C桶和E桶），如图10-7所示。但它们的令牌填充速率是不同的，C桶填充速率为CIR， P桶为PIR；两桶的容量分别为CBS和PBS（之所以用C桶和P桶表示也是基于方便描述，因为表示不同速率的参数与对应桶的容量参数相同，第一个字母对应为C，或者P）。用T_c 和T_p 表示两桶中的令牌数目，初始状态时两桶是满的，即T_c 和T_p 初始值分别等于CBS和PBS。

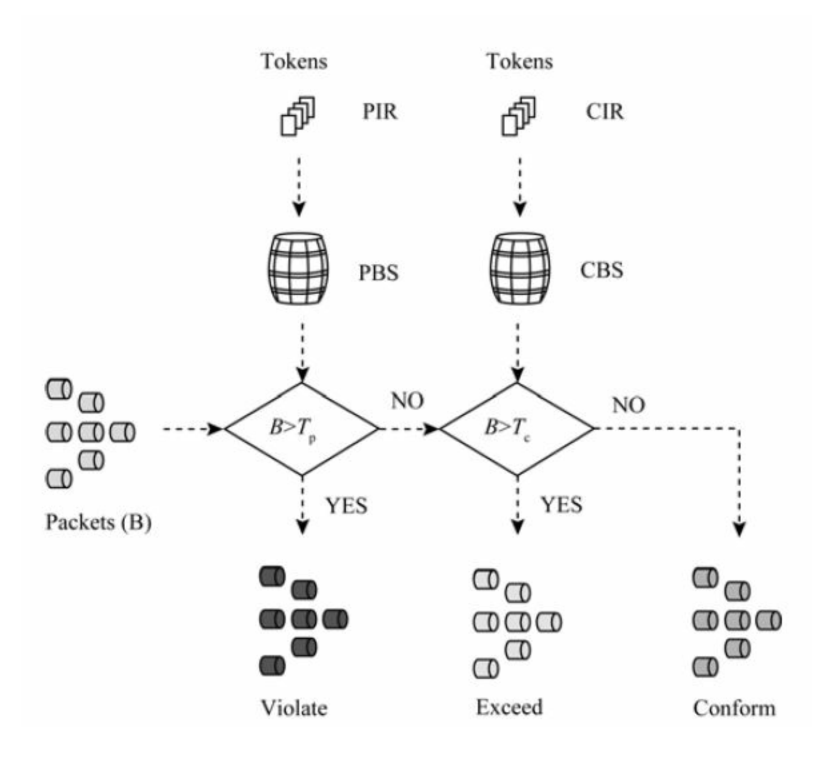


图10-7 双速双桶示意图

1．trTCM算法的4个参数

trTCM算法关注的是速率的突发，但它不像单速率三色标记算法那样把第一个桶中未使用的令牌放到第二个桶中，而是使用两个独立的令牌桶。第一个令牌桶为 PIR，大小为PBS，第二个令牌桶为CIR，大小为CBS。数据的测量是先比较PIR，再比较CIR。也就是在双速率三色标记中，首先判断的是数据发送速率是否符合规定的突发要求，而不是正常情况下的色标方法。

trTCM算法主要是根据4种流量参数来评估：CIR、CBS、峰值信息速率（PIR，Peak Information Rate），峰值突发尺寸（PBS，Peak Burst Size）。CIR和CBS参数与单速率三色算法中的含义相同，PIR就是允许的最大突发信息传输速率，，即P桶允许传输或转发报文的峰值速率，当然它的值肯定不会小于CIR的；PBS是允许的最大突发信息尺寸，表示P桶的容量，它的值也不会小于CBS。

2．trTCM算法原理

在trTCM算法中，系统按照PIR速率向P桶中填充令牌，按照CIR速率向C桶中填充令牌。

（1）当T_p <PBS时，P桶中令牌数增加，否则不增加。

（2）当T_c <CBS时，C桶中令牌数增加，否则不增加。

对于到达的报文，用B表示报文的大小。

（1）若T_p <B，报文被标记为红色。

（2）若T_c <B≤T_p ，报文被标记为黄色，且T_p 减少B。

（3）若B≤T_c ，报文被标记为绿色，且T_p 和T_c 都减少B。

3．trTCM算法中的报文着色处理

在trTCM算法中也有色盲模式和色敏模式两种。

在色盲模式下，当包速率大于PIR，此时未超过T_p +T_c 部分的包会分别从P桶和C桶中获取令牌，而且从P桶中获取令牌的部分包被标记为黄色，从C桶中获取令牌的部分包被标记为绿色，超过T_p +T_c 部分无法得

到令牌的包被标记为红色；当包速率小于PIR，而大于CIR时，包可以得到令牌，但超过T_c部分的包将从P桶中获取令牌，此时这部分包都被标记为黄色，而从C桶中获取令牌的包被标记为绿色；当包速率小于CIR时，包所需令牌数不会超过T_c，只需从C桶中获取令牌，包被标记为绿色。

在色敏模式下，如果包已被标记为红色，或者超过 T_p +T_c 部分无法得到令牌的包，被标记为红色；如果标记为黄色，或者超过 T_c 但未超过 T_p 部分包标记为黄色；如果包被标记为绿色，或者未超过T_c 部分包，被标记为绿色。

10.3.4 流量监管

"流量监管"（Traffic Policing）就是对流量进行控制，通过监督进入交换机端口的流量速率，对超出部分的流量进行"惩罚"（采用监管方式时是直接丢弃），使进入端口的流量被限制在一个合理的范围之内。例如可以限制HTTP报文不能占用超过50%的网络带宽，否则QoS流量监管功能可以选择丢弃报文，或重新配置报文的优先级。

流量监管的基本工作机制如图10-8所示，由以下三部分组成。

（1）Meter（度量器）：通过令牌桶机制对网络流量进行度量，然后向Marker（标记器）输出度量结果。

（2）Marker（标记器）：根据 Meter 的度量结果对报文进行染色，报文会被标识成green、yellow、red三种颜色。

（3）Action：根据Marker对报文的染色结果，对报文进行一些行为，行为如下。

① pass：对测量结果为"符合"的报文继续转发。

② remark + pass：修改报文内部优先级后再转发。

③ discard：对测量结果为"不符合"的报文进行丢弃。

缺省情况下，对green、yellow颜色的报文进行转发，对red报文进行丢弃。

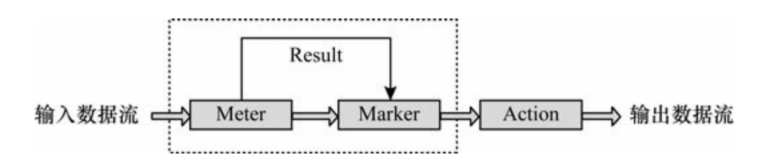


图10-8 流量监管基本工作机制示意图

当网络发生拥塞后，超出的流量将采取其他方式处理。如果处理方式为监管，那么数据包就会被丢弃。通常情况下，网络设备缺省丢弃后到的数据包而传输先到的数据包，这样的丢弃方式称为尾丢弃。也可以让网络设备在发生拥塞时，先丢弃低优先级的数据包而传输高优先级的数据包。

总体而言，经过流量监管后，如果某流量速率超过标准，设备可以选择降低报文优先级再进行转发或者直接丢弃。缺省情况下，此类报文被丢弃。如图10-9是一种经过流量监管后的流量变化示意图，超出CAR的流量均被"削"掉了。

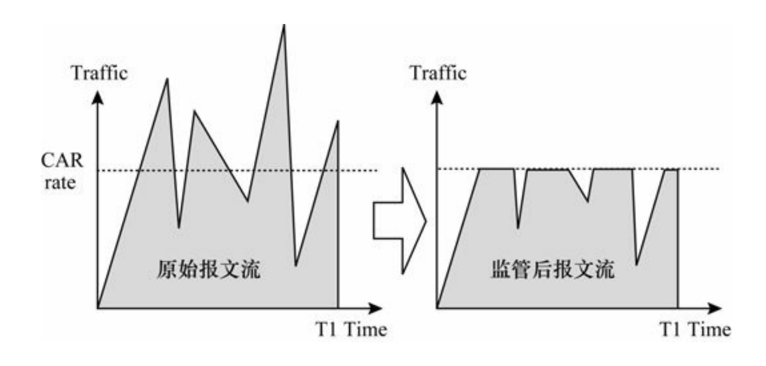


图10-9 经过浏览监管后的流量变化示意图

10.3.5 流量整形

“流量整形”为控制最大输出通信速率提供可能，以确保通信符合配置的最大传输速率规定。符合某种配置的通信可能被整形，以使它符合下游设备的通信速率需求，处理任何失配的数据传输速率。流量整形通常使用缓冲区和令牌桶来完成，当报文的发送速率过快时，首先在缓冲区进行缓存，在令牌桶的控制下再均匀地发送这些被缓冲的报文。

当下游设备的接口速率小于上游设备的端口速率或发生突发流量时，在下游设备入端口处可能出现流量拥塞的情况。此时用户可以通过在上游设备的出端口配置流量整形，将上游不规整的流量进行削峰填谷，输出一条比较平整的流量，从而解决下游设备的拥塞问题。

流量整形是一种可应用于接口、子接口或队列的流量控制技术，可以对从接口上经过的所有报文或某类报文进行速率限制。下面以接口或子接口下采用单速单桶技术的基于流的队列整形为例介绍流量整形的处理流程，其处理流程如图10-10所示。具体处理流程如下。

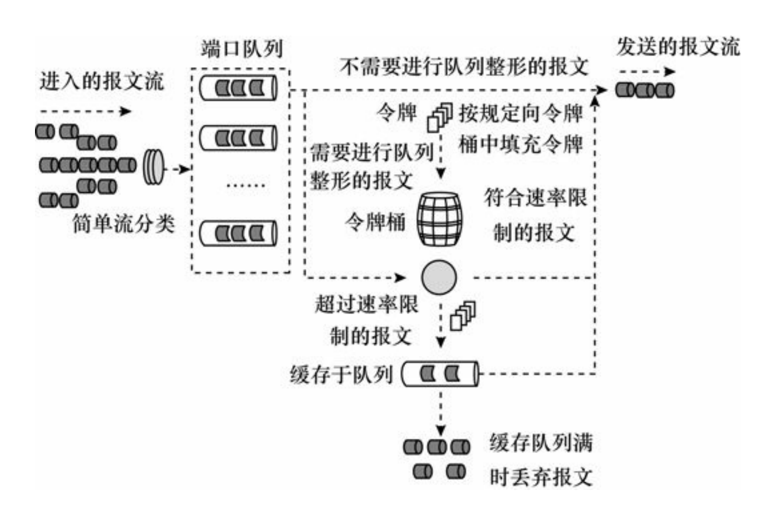


图10-10 流量整形处理流程图

（1）当报文到达设备端口时，首先对报文进行简单分类，使报文进入不同的队列。

（2）若报文进入的队列没有配置队列流量整形功能，则直接发送该队列的报文；否则，进入下一步处理。

（3）按用户设定的队列整形速率向令牌桶中存放令牌。

① 如果令牌桶中有足够的令牌可以用来发送报文，则报文直接被发送，在报文被发送的同时，令牌做

相应的减少。

② 如果令牌桶中没有足够的令牌，则将报文放入缓存队列，如果报文放入缓存队列时，缓存队列已满，则丢弃报文。

（4）缓存队列中有报文的时候，系统按一定的周期从缓存队列中取出报文进行发送，每次发送都会与令牌桶中的令牌数作比较，直到令牌桶中的令牌数减少到缓存队列中的报文不能再发送或缓存队列中的报文全部发送完毕为止。

说明

经过以上队列整形后，如果该接口或子接口同时配置了基于端口的流量整形功能， 则系统还要逐级按照子接口整形速率、接口整形速率对报文流进行速率控制。其处理流程与上述流程相似，但不需要步骤1和2。

流量整形和流量监管都是作用于网络边缘，对进入设备端口的流量进行的一种处理方式。它们的主要区别在于：流量监管直接丢弃不符合速率要求的报文，丢弃的报文比较多，可能引发重传；而流量整形是将不符合速率要求的报文先行缓存，当令牌桶有足够的令牌时再均匀地向外发送这些被缓存的报文，较少丢弃报文，但引入时延和抖动，需要较多的缓冲资源缓存报文。所以这两种功能的应用领域也不尽相同，流量监管适用于对丢弃率不敏感，而对时延和抖动比较敏感的网络应用，如一些普通的话音和视频通信；流量整形适用于对时延和抖动不敏感的网络应用，如数据传输、WWW访问等。

10.4 拥塞避免和拥塞管理

当网络间歇性地出现拥塞，且时延敏感业务要求得到比非时延敏感业务更高质量的 QoS服务时，需要进行拥塞管理；如果配置拥塞管理后仍然出现拥塞，则需要增加带宽。拥塞避免（Congestion Avoidance）是指通过监视网络资源（如队列或内存缓冲区）的使用情况，在拥塞发生或有加剧的趋势时主动丢弃报文，通过调整网络的流量来解除网络过载的一种流控机制。

为了解决网络拥塞，可以通过拥塞避免 在网络出现拥塞时主动丢弃一些报文，解除网络过载；为了使用户得到更好的服务质量，可以通过拥塞管理 对关键业务优先调度，使得这些业务得到更高的QoS服务。

10.4.1 拥塞避免

华为S系列交换机主要支持以下两种拥塞避免功能。

1．尾部丢弃

传统的丢包策略采用尾部丢弃的方法，且是同等地对待所有报文，不区分报文的服务等级。这时在拥塞发生期间，队列尾部的数据报文将被丢弃，直到拥塞解除。但这种丢弃策略会引发TCP全局同步现象。所谓TCP全局同步现象，是指当多个队列同时丢弃多个TCP连接报文时，将造成多个TCP连接同时进入拥塞避免和慢启动状态，以降低并调整流量；而后这几个TCP连接又会在某个时刻同时出现流量高峰。如此反复，使网络流量忽大忽小，影响链路利用率。

2．RED/SRED/WRED

在华为S系列交换机中，对RED（Random Early Detection，随机早期检测）、SRED （Simple Random Early Detection，简单随机早期检测）和WRED（Weighted Random Early Detection，权重随机早期检测）三种拥塞避免技术有不同的支持。

RED 技术是通过随机地丢弃报文让多个 TCP 连接不同时降低发送速率，从而避免了TCP的全局同步现象。在RED技术的算法中，为每个队列的长度都设定了阈值的上、下限值，并有以下规定。

（1）当队列的长度小于阈值下限时，不丢弃报文。

（2）当队列的长度大于阈值上限时，丢弃所有新收到的报文。

（3）当队列的长度在阈值上限和阈值下限之间时，开始随机丢弃到来的报文。方法是为每个到来的报文赋予一个随机数，并用该随机数与当前队列的丢弃概率比较，如果大于丢弃概率则报文被丢弃。队列越长，报文被丢弃的概率越高。

SRED技术是在RED技术基础上诞生的。在接口出队列上，SRED会根据报文的优先级（而不是随机选择要丢弃的报文）将其区分为红色、黄色，并分别为红色和黄色的报文设定起始丢包点和丢包率，然后通过按照一定的丢弃概率主动丢弃队列中的红色甚至黄色报文，从而调整从接口输出的流量速率。

WRED技术也是在RED技术基础上改进的，与SRED一样也是基于报文优先级来选择丢弃报文的，但WRED同时还可为相同颜色的不同报文设置不同的丢弃权重，以实现更加灵活的报文丢弃策略，使高优先级报文被丢弃的概率相对较小。

10.4.2 拥塞管理

拥塞管理一般采用队列调度技术，使用不同的调度算法来发送队列中的报文流。根 据排队和调度策略的不同，设备LAN接口上的队列调度技术分为PQ、DRR、PQ+DRR、WRR、PQ+WRR；WAN接口上的队列调度技术分为PQ、WFQ、PQ+WFQ和CBQ。下面具体介绍这些调度方法。

1．PQ调度

PQ（Priority Queueing，优先队列）调度是针对于关键业务类型应用而设计的队列机制。PQ 调度算法维护一个优先级递减的队列系列，并且只有当更高优先级的所有队列为空时才服务低优先级的队列。这样，将关键业务的分组放入较高优先级的队列，将非关键业务（如E-mail）的分组放入较低优先级的队列，可以保证关键业务的分组被优先传送，非关键业务的分组在处理关键业务数据的空闲间隙被传送。

如图10-11所示，Queue7比Queue6具有更高的优先权，Queue6比Queue5具有更高的优先权，依此类推。只要链路能够传输分组，Queue7 尽可能快地被服务。只有当Queue7为空，调度器才考虑Queue6。当Queue6有分组等待传输且Queue7为空时，Queue6以链路速率接受类似的服务。当Queue7和Queue6为空时，Queue5以链路速率接收服务，依此类推。

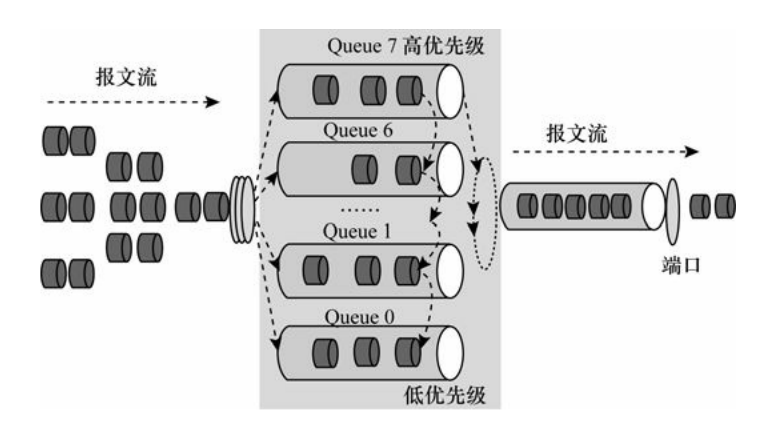


图10-11 PQ调度示意图

PQ调度算法对低时延业务非常有用。假定数据流X在每一个节点都被映射到最高优先级队列，那么当数据流X的分组到达时，则分组将得到优先服务。然而PQ调度机制会使低优先级队列中的报文由于得不到服务而“饿死”。例如，如果映射到Queue7的数据流在一段时间内以100%的输出链路速率到达，调度器将从

不为Queue6及以下的队列服务。所以在采用 PQ 调度方式时，应将延迟敏感的关键业务放入高优先级队列，将非关键业务放入低优先级队列，从而确保关键业务被优先发送。

2．WRR调度

WRR（Weight Round Robin，加权循环调度）是在RR（Round Robin，循环调度）的基础上演变而来的。它可在队列之间进行轮流调度，根据每个队列的权重来调度各队列中的报文流。实际上，RR调度相当于权值为1（即每个队列在调度一次后都重新开始新的一轮调度）的WRR调度。WRR队列示意图如图10-12所示。

在进行WRR调度时，设备根据每个队列的权值进行轮循调度。调度一轮权值减一，权值减到零的队列不参加调度，当所有队列的权值减到0时，开始下一轮的调度。例如， 用户根据需要为接口上8个队列指定的权值分别为4、2、5、3、6、4、2和1，按照WRR方式进行调度的结果请参见表10-6。

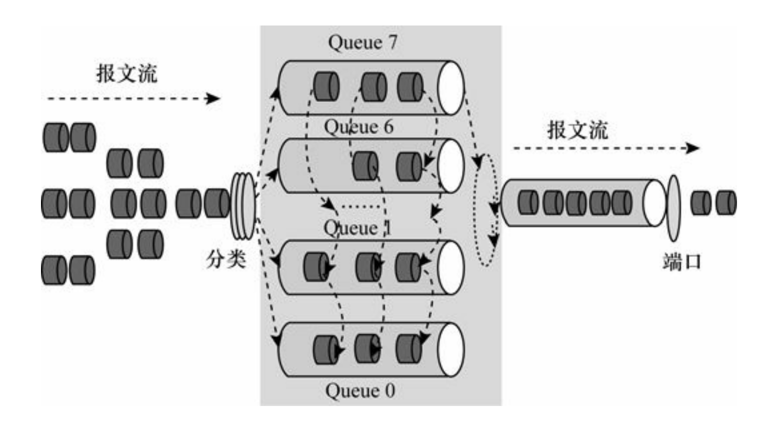


图10-12 WRR调度示意图

表10-6 WRR调度示例的调度结果

队列索引	Q7	Q6	Q5	Q4	Q3	Q2	Q1	Q0
队列权值	4	2	5	3	6	4	2	1
参加第 1 轮调度的队列	Q7	Q6	Q5	Q4	Q3	Q2	Q1	Q0
参加第 2 轮调度的队列	Q7	Q6	Q5	Q4	Q3	Q2	Q1	—
参加第 3 轮调度的队列	Q7	—	Q5	Q4	Q3	Q2	—	—
参加第 4 轮调度的队列	Q7	—	Q5	—	Q3	Q2	—	—
参加第 5 轮调度的队列	—	—	Q5	—	Q3	—	—	—
参加第 6 轮调度的队列	—	—	—	—	Q3	—	—	—
参加第 7 轮调度的队列	Q7	Q6	Q5	Q4	Q3	Q2	Q1	Q0
参加第 8 轮调度的队列	Q7	Q6	Q5	Q4	Q3	Q2	Q1	—
参加第 9 轮调度的队列	Q7	—	Q5	Q4	Q3	Q2	—	—
参加第 10 轮调度的队列	Q7	—	—	Q4	Q3	Q2	—	—
参加第 11 轮调度的队列	—	—	Q5	—	Q3	—	—	—
参加第 12 轮调度的队列	—	—	—	—	Q3	—	—	—

从表10-6可以看出，各队列中的报文流被调度的次数与该队列的权值成正比，权值越大被调度的次数相对越多。由于WRR调度是以报文为单位，因此每个队列没有固定的带宽，同等调度机会下大尺寸报文获得的实际带宽要大于小尺寸报文获得的带宽，避免了采用 PQ 调度时低优先级队列中的报文可能长时间得不到服务的缺点。另外， WRR调度中虽然多个队列的调度是轮循进行的，但对每个队列不是固定地分配服务时间片——如果某个队列为空，那么马上换到下一个队列调度，这样带宽资源可以得到充分的利用。

但WRR调度有两个缺点。

（1）WRR 调度按照报文个数进行调度，而用户一般关心的是带宽。当每个队列的平均报文长度相等或已知时，通过配置 WRR 权重，用户能够获得想要的带宽；但是， 当队列的平均报文长度变化时，用户就不能通过配置WRR权重获取想要的带宽。

（2）低延时需求业务（如话音）得不到及时调度。

3．DRR调度

DRR（Deficit Round Robin，差额循环调度）调度同样也是RR（循环调度）法的扩展。相对于WRR而言，DRR调度解决了WRR调度中只关心报文，同等调度机会下大尺寸报文获得的实际带宽要大于小尺寸报文获得的带宽的问题。DRR调度通过调度过程中考虑了包长的因素，从而达到调度的速率公平性。

在DRR调度中，Deficit表示队列的带宽赤字，初始值为0。每次调度前，系统按权重为各队列分配带宽，计算Deficit值，如果队列的Deficit值大于0，则参与此轮调度，发送一个报文，并根据所发送报文的长度计算调度后Deficit值，作为下一轮调度的依据；如果队列的Deficit值小于0，则不参与此轮调度，当前Deficit值作为下一轮调度的依据。

假设用户配置各队列权重为40、30、20、10、40、30、20、10（依次对应Q7、Q6、Q5、Q4、Q3、Q2、Q1、Q0）。在调度初始时，队列Q7、Q6、Q5、Q4、Q3、Q2、Q1、Q0依次能够获取20%、15%、10%、5%、20%、15%、10%、5%的带宽。下面以Q7、Q6为例，简要描述DRR队列调度的实现过程（假设Q7队列获取400bit/s的带宽，Q6队列获取300bit/s的带宽）。

第**1**轮调度

Deficit[7] [1] = 0（为Q7初始Deficit值）+400（为Q7队列带宽）= 400（为Q7队列第一轮调度时的Deficit值），Deficit[6] [1] = 0（为Q6初始Deficit值）+300（为Q6队列带宽）= 300（为Q6队列第 1轮调度时的Deficit值）。

假设这时从Q7队列取出一个900bytes的报文发送，从Q6队列取出一个400bytes的报文发送。发送后，Deficit[7] [1] 的 Deficit 值=400–900（400 为 Q7 队列的带宽，900为Q7队列此次发送的数据字节大小）=－500（为Q7队列第1轮调度后的Deficit值）， Deficit[6] [1] 的Deficit值= 300–400（300为Q6队列的带宽，400为Q6队列此次发送的数据字节大小）=－100（为Q6队列第1轮调度后的Deficit值）。

第**2**轮调度

Deficit[7] [2] =－500（为Q7第 1轮调度后的Deficit值）+400（为Q7队列的带宽）=－100（为Q7第 2轮调度时可用的Deficit值），同理，Deficit[6] [2] =－100+300 = 200。因为Q7队列第2轮调度Deficit值小于0，所以此轮不参与调度，而Q6队列第2轮调度Deficit值大于0，所以仍可以从Q6队列取出一个报文（假设为300bytes）进行发送。发送后，Deficit[6] [2] = 200–300 =－100（为Q6队列第 2轮调度后的Deficit值）。

第**3**轮调度

Deficit[7] [3] =－100+400 = 300，Deficit[6] [3] =－100+300 = 200。因为两个队列此时的Deficit值均大0，所以均可以发送数据。假设此时从Q7队列取出一个600bytes的报文发送，从Q6队列取出一个 400bytes的报文发送。发送后，Deficit[7] [3] = 300–600 =－300，Deficit[6] [3] = 200–500 =－300。这时再进行第 4轮调度时，可根据前面介绍的方法计算第4轮调度时两队列的Deficit值分别为100和0，所以Q7队列可以继续发送报文，而Q6队列不允许发送报文。如此循环调度，最终Q7、Q6队列获取的带宽将分别占总带宽的20%、15%。因此，通过DRR调度方法时用户能够通过设置权重获取想要的 带宽。但DRR调度仍然没有解决WRR调度中低延时需求业务得不到及时调度的问题。

DRR调度避免了采用PQ调度时低优先级队列中的报文可能长时间得不到服务的缺点，也避免了各队列报文长度不等或变化较大时，WRR 调度不能按配置比例分配带宽资源的缺点。但是，DRR调度也具有低延时需求业务（如话音）得不到及时调度的缺点。

4．WFQ调度

FQ（Fair Queuing，公平队列）的目的是尽可能公平地分享网络资源，使所有流的延迟和抖动达到最优。即不同的队列获得公平的调度机会，从总体上均衡各个流的延迟；短报文和长报文也可获得公平的调度：如果不同队列间同时存在多个长报文和短报文等待发送，让短报文优先获得调度，从而在总体上减少各个流的报文间的抖动。

WFQ调度在报文入队列之前，先对流量进行分类，有两种分类方式。

（1）按流的“会话”信息分类。根据报文的协议类型、源和目的TCP或UDP端口号、源和目的IP地址、ToS域中的优先级位等自动进行流分类，并且尽可能多地提供队列，以将每个流均匀地放入不同队列中，从而在总体上均衡各个流的延迟。在出队的时候，WFQ按流的优先级来分配每个流应占有带宽。优先级的数值越小，所得的带宽越少。优先级的数值越大，所得的带宽越多。

（2）按优先级分类。通过优先级映射把流量标记为本地优先级，每个本地优先级对应一个队列号。每个接口预分配4个或8个队列，报文根据队列号进入队列。缺省情况，队列的 WFQ 权重相同，流量平均分配接口带宽。用户可以通过配置修改权重，高优先权和低优先权按权重比例分配带宽。

以上整个WFQ调度原理如图10-13所示。以端口有8个输出队列为例，WRR可为每个队列配置一个加权值（依次为 w7、w6、w5、w4、w3、w2、w1、w0），加权值表示获取资源的比重。例如：一个100MB的端口，配置它的WRR队列调度算法的加权值为50、50、30、30、10、10、10、10（依次对应w7、w6、w5、w4、w3、w2、w1、w0），这样可以保证最低优先级队列至少获得5Mbit/s带宽，避免了采用PQ调度时低优先级队列中的报文可能长时间得不到服务的缺点。从统计上，WFQ使高优先权的报文获得优先调度的机会多于低优先权的报文。

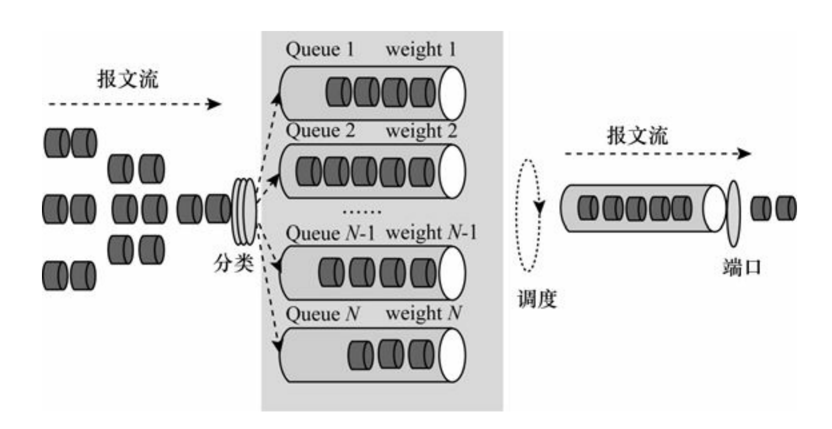


图10-13 WFQ调度示意图

5．PQ+WRR调度

PQ调度和WRR调度各有优缺点，为了克服单纯采用PQ调度或WRR调度时的缺 点，PQ+WRR调度以发挥两种调度的各自优势，不仅可以通过WRR调度可以让低优先级队列中的报文也能及时获得带宽，而且通过 PQ 调度可以保证了低延时需求的业务能优先得到调度。

在设备上，用户可以配置队列的WRR参数，根据配置将接口上的8个队列分为两组，一组（例如Queue7、Queue6、Queue5）采用PQ调度，另一组（例如Queue4、Queue3、Queue2、Queue1和Queue0队列）采用WRR调度。设备上只有LAN侧接口支持PQ+WRR调度。

PQ+WRR 调度示意图如图 10-14 所示。在调度时，设备首先按照 PQ 方式调度Queue7、Queue6、Queue5队列中的报文流，只有这些队列中的报文流全部调度完毕后，才开始以WRR方式循环调度其他队列

中的报文流。Queue4、Queue3、Queue2、Queue1和Queue0队列包含自己的权值。重要的协议报文和有低延时需求的业务报文应放入采用PQ调度的队列中，得到优先调度的机会，其余报文放入以WRR方式调度的各队列中。

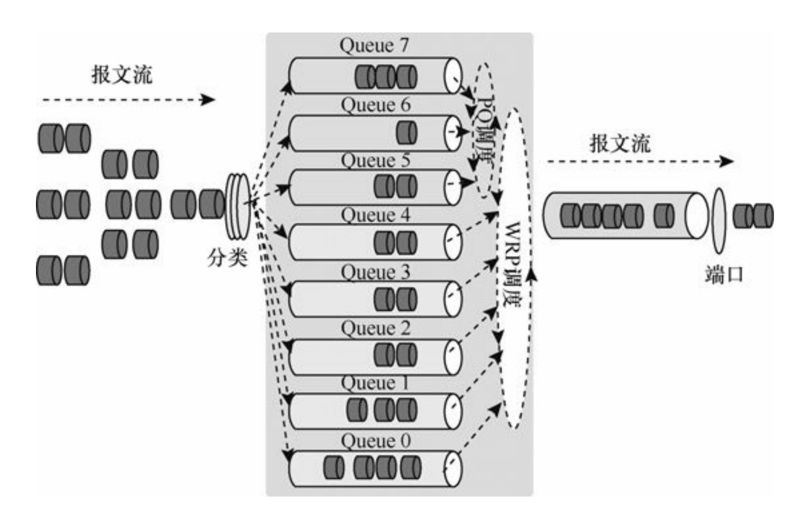


图10-14 PQ+WRR混合调度示意图

6．PQ+DRR调度

与PQ+WRR相似，其集合了PQ调度和DRR调度，各有优缺点。单纯采用PQ调度时，低优先级队列中的报文流长期得不到带宽，而单纯采用 DRR 调度时低延时需求业务（如话音）得不到优先调度，如果将两种调度方式结合起来形成PQ+DRR调度，不仅能发挥两种调度的优势，而且能克服两种调度各自的缺点。

在PQ+DRR调度中，设备接口上的8个队列被分为两组，用户可以指定其中的某几组队列进行PQ调度，其他队列进行DRR调度。如图10-15所示，在调度时设备首先按照PQ方式优先调度Queue7、Queue6和Queue5队列中的报文流，只有这些队列中的报文流全部调度完毕后，才开始以 DRR 方式调度 Queue4、Queue3、Queue2、Queue1 和 Queue0 队列中的报文流。其中，Queue4、Queue3、Queue2、Queue1和Queue0队列包含自己的权值。重要的协议报文以及有低延时需求的业务报文应放入需要进行PQ调度的队列中，得到优先调度的机会，其他报文放入以DRR方式调度的各队列中。

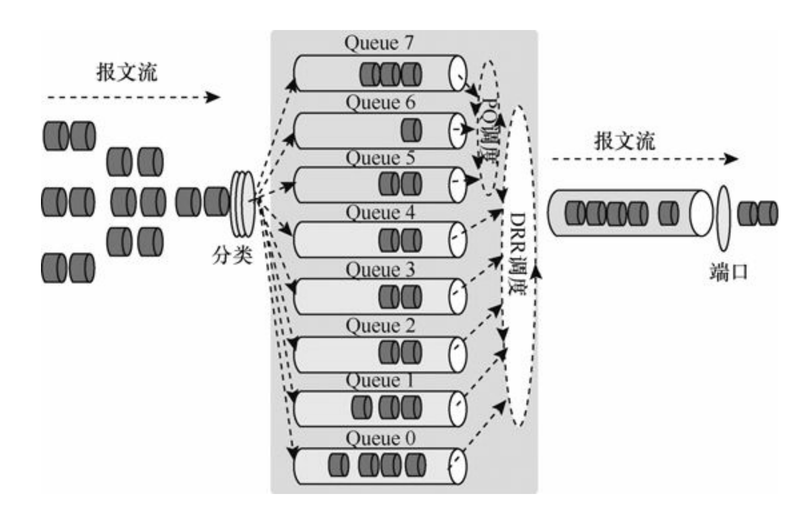


图10-15 PQ+WRR调度示意图

7．PQ+WFQ调度

与PQ+WRR相似，PQ+WFQ调度方式集合了PQ调度和WFQ调度，各有优缺点。单纯采用PQ调度时，低优先级队列中的报文流长期得不到带宽，而单纯采用WFQ调度时低延时需求业务（如话音）得不到优先调度，如果将两种调度方式结合起来形成PQ+WFQ调度，不仅能发挥两种调度的优势，而且能克服两种调度各自的缺点。

在 PQ+WFQ 调度中，设备接口上的 8 个队列也被分为两组，用户可以指定其中的某几组队列进行PQ调度，其他队列进行WFQ调度。只有WAN侧接口支持PQ+WFQ调度。如图10-16所示，在调度时，设备首先按照PQ方式优先调度Queue7、Queue6和Queue5队列中的报文流，只有这些队列中的报文流全部调度完毕后，才开始以WFQ方式调度Queue4、Queue3、Queue2、Queue1和Queue0队列中的报文流。其中，Queue4、Queue3、Queue2、Queue1和Queue0队列包含自己的权值。重要的协议报文以及有低延时需求的业务报文应放入需要进行 PQ 调度的队列中，得到优先调度的机会，其他报文放入以WFQ方式调度的各队列中。

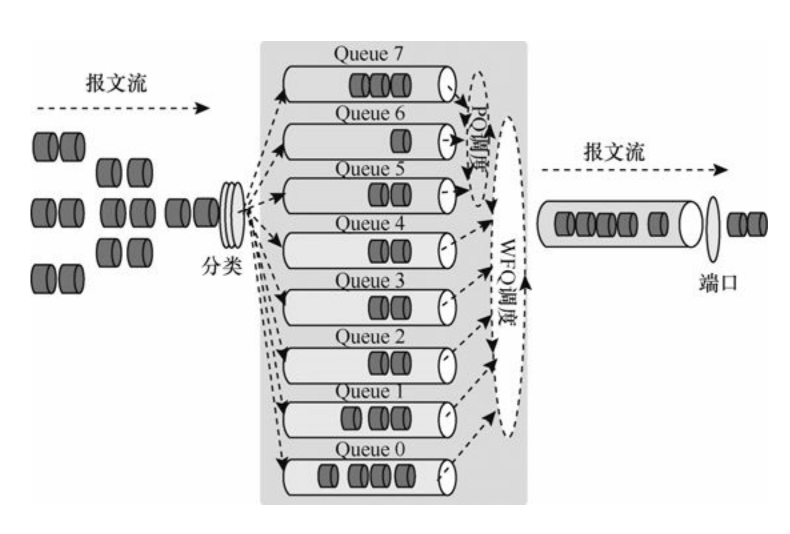


图10-16 PQ+WFQ调度示意图

8．CBQ调度

CBQ（Class-based Queueing，基于类的加权公平队列）是对WFQ功能的扩展，为用户提供了定义类的支持。CBQ首先根据IP优先级或者DSCP优先级、输入接口、IP报文的五元组等规则对报文进行分类，然后让不同类别的报文进入不同的队列。对于不匹配任何类别的报文，送入系统定义的缺省类。如图10-17所示，CBQ提供以下三类队列。

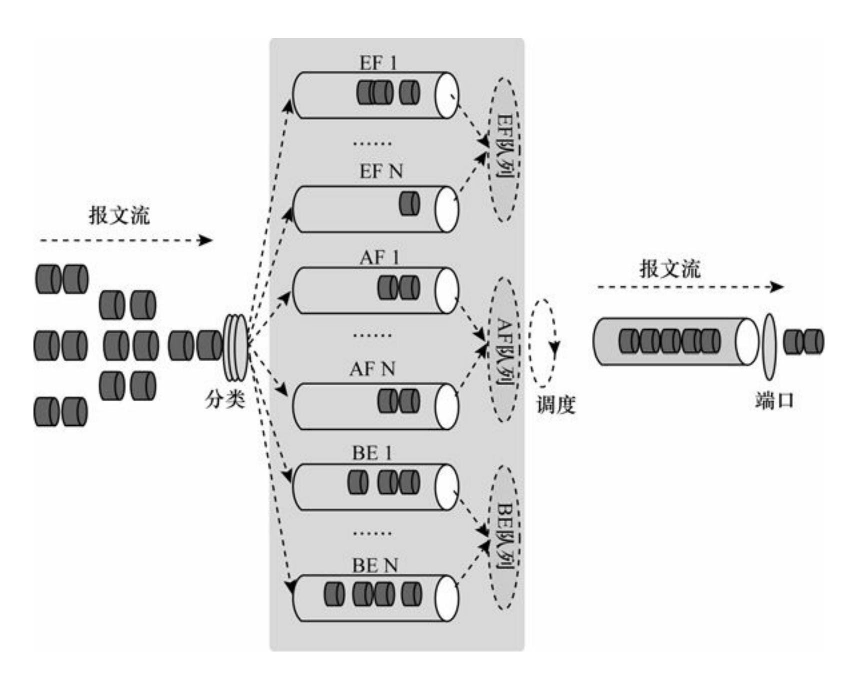


图10-17 CBQ调度示意图

（1）EF（加速转发）队列：满足低时延业务

EF队列是具有高优先级的队列，一个或多个类的报文可以被设定进入EF队列，不同类别的报文可设定占用不同的带宽。在调度出队的时候，若 EF 队列中有报文，会优先得到调度，以保证其获得低时延。当接口发生拥塞时，EF队列的报文会优先发送，但为了防止低优先级队列（AF、BE队列）得不到调度，EF队列以设置的带宽限速。当接口不拥塞时，EF队列可以占用AF、BE的空闲带宽。这样，属于EF队列的报文既可以获得空闲的带宽，又不会占用超出规定的带宽，保护了其他报文的应得带宽。

设备除了提供普通的EF队列，还支持一种特殊的EF队列——LLQ（低时延队列）。两种队列都采用绝对优先调度，但是LLQ队列使用流量监管实现，不论接口是否拥塞，流量都不会超过设置的带宽，LLQ队列不缓存报文，可以将报文被发送的时延降低为最低限度。这为对时延敏感的应用（如VoIP业务）提供了良好的服务质量保证。

（2）AF（确保转发）队列：满足需要带宽保证的关键数据业务

每个 AF 队列分别对应一类报文，用户可以设定每类报文占用的带宽。在系统调度报文出队的时候，按用户为各类报文设定的带宽将报文出队发送，可以实现各个类的队列的公平调度。当接口有剩余带宽时，AF 队列按照权重分享剩余带宽。同时，在接口拥塞的时候，仍然能保证各类报文得到用户设定的最小带宽。

对于 AF 队列，当队列的长度达到队列的最大长度时，缺省采用尾丢弃的策略，但 用户还可以选择用WRED丢弃策略。

（3）BE队列（尽力而为）：满足不需要严格QoS保证的尽力发送业务

当报文不匹配用户设定的所有类别时，报文被送入系统定义的缺省类。虽然允许为缺省类配置AF队列，并配置带宽，但是更多的情况是为缺省类配置BE队列。BE队列使用WFQ调度，使所有进入缺省类的报文进行基于流的队列调度。

对于 BE 队列，当队列的长度达到队列的最大长度时，缺省采用尾丢弃的策略，但用户还可以选择用WRED丢弃策略。

10.5 流策略

流策略是指按照某种规则对流量进行分类，并对同种类型的流量关联某种行为，形成某种策略。将该策略应用后可实现流量监管、重新标记优先级、重定向等功能。

流策略包含3个要素：流分类、流行为和流策略，这也是配置QoS流策略的前三大基本任务（最后一个任务就是应用流策略）。下面分别予以介绍。

1．流分类

流分类采用一定的规则识别符合某类特征的报文，从而把具有某类共同特征的报文划分为一类，它是有区别地进行服务的前提和基础。

用户可以通过定义一系列的规则来对报文进行分类，同时也可以指定规则之间的关系。

（1）and：报文只有匹配了类中的所有的规则，设备才认为报文属于此类。当流分类中有ACL规则时，报文必须匹配其中一条ACL规则以及所有非ACL规则才属于该类。当流分类中没有ACL规则时，则报文必须匹配所有非ACL规则才属于该类。

（2）or：报文只要匹配了类中的一个规则，设备就认为报文属于此类。

可以使用的流分类规则如表10-7所示。

表10-7 流分类的分类规则

二层分类规则	三层分类规则	其他分类规则
• 目的 MAC 地址 • 源 MAC 地址 • VLAN 报文外层 Tag 的 ID 信息 • VLAN 报文外层 Tag 的 802.1p 优先级 • VLAN 报文内层 Tag 的 ID 信息 • VLAN 报文内层 Tag 的 802.1p 优先级 • 基于二层封装的协议字段 • FR 报文中的 DE 标志位 • FR 报文中的 DLCI 信息 • ATM 报文中的 PVC 信息 • ACL 4 000～4 999（二层 ACL）	• IP 报文的 DSCP 优先级 • IP 报文的 IP 优先级 • IP 协议类型（IPv4 协议或IPv6 协议） • RTP 端口号 • TCP 报文的 TCP-Flag • ACL 2 000～3 999（基本ACL） • ACL6 2 000～3 999（高级ACL）	• 入接口 • 出接口 • ACL 5 000～5 999（用户自定义 ACL） • SAC（Smart Application Control，灵活应用控制）

2．流行为

流行为用来定义针对某类报文所做的QoS行为。进行流分类是为了有区别地提供服 务，它必须与某种流量控制或资源分配的流行为关联起来才有意义。

针对流分类可实施的流行为包括报文过滤、重标记、重定向、流量监管、流量整形、流镜像、队列调度、流量统计、绑定子流策略、禁止URPF检查、封装外层VLAN标签和禁止MAC地址学习。具体介绍如下。

（1）报文过滤。报文过滤是最简单的流控行为。通过对报文的允许或禁止行为处理，控制网络流量，实现动态的防火墙报文过滤功能。

（2）重标记。重标记是指将报文的优先级字段进行重新设置。在不同的网络中报文使用不同的优先级字段，例如VLAN网络使用802.1p，IP网络使用ToS，MPLS网络使用EXP。因此需要设备可以针对不同的网络对报文的优先级进行重标记。

通常是需要在网络的边界节点设备上对进入的报文进行优先级重标记，网络内部的节点设备按照边界节点所标记的优先级提供相应等级的QoS服务，或者按自己的标准重新进行标记。

（3）重定向。重定向是指将报文不按原始的目的地址进行路由转发，而是将报文重定向转发到CPU、指定接口、指定的下一跳地址或下一跳标签LSP。

通过重定向可以实现策略路由，这也是QoS策略的一种主要应用。这种策略路由是静态的，当配置中的下一跳不可用时，系统将按原来的转发路径转发报文。

（4）流量监管。流量监管是一种通过对流量规格的监督，来限制流量及其资源使用的流控行为。为了避免用户不断突发的业务数据造成网络拥挤，可以通过流量监管，控制某些匹配分类规则的流的规格，对于超过规格的流量，可以采取丢弃、重标记颜色、重标记优先级或其他的QoS措施，以更好地利用网络资源。

（5）流量整形。流量整形也是一种通过对流量规格的监督，来限制流量及其资源使用的流控行为。它是一种主动调整流的输出速率的流控措施，通常是为了使流量适配下游设备可供给的网络资源，避免不必要的报文丢弃和拥塞。流量整形通过限制流出某一网络的某一连接的流量，使这类报文以比较均匀的速度向外发送。

（6）流镜像。即将指定的数据包复制到用户指定的目的观察端口，以进行网络流量监控和故障分析与排除。有关端口镜像配置可参见本书第14章。

（7）队列调度。通常用来对某类的流量进行拥塞管理和拥塞避免，使不同类型的业务流进入不同优先级的队列，以获取不同等级的QoS转发服务。

（8）流量统计。流量统计用于统计指定业务流的数据报文。它统计的是匹配流分类的报文中通过和丢弃的报文数量和字节数。但流量统计本身不是QoS控制措施，但可以和其他QoS行为组合使用，以提高网络和报文的安全性。

（9）绑定子流策略。绑定子流策略是指为流策略中的流行为绑定一个子流策略，实现流策略嵌套。使用流策略嵌套时，对于匹配流分类的某一类报文，除了执行父策略中定义的行为外，还由子策略对该类流量进行再次分类，执行子策略中定义的行为，实现了更为精细化的HQoS（高级QoS）服务。

（10）禁止URPF检查。禁止URPF（Unicast Reverse Path Forward，单播反向路径转发）检查是指设备对符合流分类规则的报文不进行逆向地址检查。配置接口的URPF检查功能后，设备对进入接口的所有报文都进行 URPF 检查，丢弃源地址对应的接口与入接口不一致的报文。此时，如果要保证某类特定的报文不被丢弃，比如设备相信从 某个服务器过来的所有报文，不对其进行 URPF 检查，可以配置对指定流禁止URPF检查功能。

（11）封装外层VLAN标签。封装外层VLAN标签是指对符合流分类规则的报文创建外层 VLAN 标签。当下游设备根据指定的外层 VLAN 标签提供差分服务时，可以在本设备上为指定流分类的报文配置封装外层VLAN 标签，以便于下游设备进行识别。

（12）禁止MAC地址学习。禁止MAC地址学习是指设备不再学习符合流分类规则的报文的 MAC 地址。在网络比较稳定，报文的 MAC 地址相对固定的情况下，为了节省MAC地址表项的开支，提高设备的运行效率，可以去使能MAC学习功能。

（13）Netstream 统计采样。对匹配流分类的流量使用 NetStream 统计采样的方法，通过设定适当的统计采样方式及采样间隔，只对匹配流分类的IPv4报文进行流信息统计分析，收集到的统计信息可以基本反映该流的流量状况，同时也可以降低使能NetStream功能对设备性能的影响。

3．流策略

流策略是将流分类和流行为绑定后形成的完整的策略。通过将流策略应用到接口、全局、单板或者VLAN，实现了针对不同业务的差分服务。

第11章　QoS配置与管理

在QoS的各项功能配置中，最基本的是报文优先级（包括二层的802.1p优先级，三层的DSCP优先级和IP优先级）与内部优先级（即端口优先级），以及PHB（逐跳行为）/颜色之间的映射配置（这是本章的重点与难点），因为在流量监管、流量整形、拥塞避免、拥塞管理和QoS流策略中大多数情况下要使用到报文优先级的映射。而在QoS流策略的配置中，整个配置思路非常清晰，就是包括定义流分类、定义流行为、创建流策略、应用流策略四大配置任务。

本章将详细介绍以上各项QoS功能和QoS流策略配置方法，并给出大量实用的配置示例。

11.1 QoS优先级映射配置与管理

优先级映射是整个 QoS 应用配置中的基础配置，许多 QoS 应用配置中都要依据报文中映射后的优先级类型和值对报文进行处理，如本章后面将要介绍的流量监管、流量整形、拥塞避免和拥塞管理、QoS流分类等。在整个华为S系列交换机中，不同系列交换机所支持的优先级映射特性及配置方法不完全一样，下面将分别介绍。

11.1.1 S2700SI/2700EI/2710SI优先级映射配置与管理

S2700SI/2700EI/2710SI系列交换机仅支持根据**VLAN**报文中的**802.1p**优先级进行优先级映射 。对于带VLAN 标签的报文，入方向根据报文携带的 802.1p 优先级按照缺省的映射关系将802.1p优先级映射为内部优先级，参见表10-6；对于不带VLAN标签的报文，设备将使用端口的缺省802.1p优先级（可配置），将端口缺省的802.1p优先级映射到内部优先级。

1．基本配置任务

在整个优先级映射配置中包括以下3项配置任务。

（1）配置端口信任的报文优先级，即优先级信任模式的配置。

（2）（可选）配置端口优先级（仅用于不带**VLAN**标签的报文优先级映射 ）。

（3）（可选）配置内部优先级和队列之间的映射关系（仅当需要改变缺省的映射关系时配置 ）。

2．具体配置步骤

以上3项配置任务的具体配置步骤如表11-1所示。对于那些在接口视图下进行的配置，如果需要在多个端口下配置相同的配置，则可选择在对应端口组视图下进行配置，以减少配置工作量。

表11-1 S2700SI/2700EI/2710SI系列优先级映射配置步骤

配置任务	步骤	命令	说明
配置端口信任的报文优先级	1	**system-view** 例如：<HUAWEI> **system-view**	进入系统视图
	2	**interface** *interface-type interface-number* 例如：[HUAWEI] **interface ethernet 0/0/1**	键入要配置端口信任的报文优先级的接口，进入接口视图
	3	**trust 8021p** 例如：[HUAWEI-Ethernet0/0/1] **trust 8021p**	配置端口信任的报文优先级为 802.1p 优先级，使得端口根据 802.1p 优先级对应的映射关系进行映射处理
		trust 8021p 例如：[HUAWEI-Ethernet0/0/1] **trust 8021p**	缺省情况下，不信任任何报文优先级，可用 **undo trust** 命令取消端口的信任优先级配置 当 S2700EI 子系列交换机不配置信任任何报文优先级时，报文都进入队列 0，且报文的 802.1p 值被设置为 0；当 S2700SI 子系列交换机不信任任何报文优先级时，报文都进入队列 0，但报文的 802.1p 值不发生变化

（续表）

配置任务	步骤	命令	说明
（可选）配置端口优先级	4	**port priority** *priority-value* 例如：[HUAWEI-Ethernet0/0/1] **port priority** 1	配置端口的缺省 802.1p 优先级，取值范围为 0～7 的整数，值越大优先级越高。如果当前接口已加入 Eth-Trunk，则本命令不可用 **端口优先级仅在收到不带 VLAN 标签的报文时有用**，此时会在设备内部转发时需要为这些不带 VLAN 标签的报文添加端口的缺省 802.1p 优先级；如果此时端口已通过上一步配置了信任 802.1p 优先级，会使用端口的缺省 802.1p 优先级查找 802.1p 优先级到内部优先级映射表，为这些不带 VLAN 标签的报文标记内部优先级 缺省情况下，接口为不带 VLAN 标签的报文添加的缺省 802.1p 优先级值为 0，可使用 **undo port priority** 命令设置接口为不带 VLAN 标签的报文添加的缺省 802.1p 优先级值 0
（可选）配置内部优先级和队列之间的映射关系	5	**quit** 例如：[HUAWEI-Ethernet0/0/1] **quit**	退出接口视图，返回系统视图
	6	**qos local-precedence-queue-map** *local-precedence queue-index* 例如：[HUAWEI] **qos local-precedence-queue-map** af3 2	配置内部优先级和入队列之间的映射关系（**如采用表 10-7 右部分的缺省映射关系，则不进行本项配置**）。命令中的参数说明如下。 （1）*local-precedence*：指定映射关系中的本地优先级名称，可以是 **af1**、**af2**、**af3**、**af4**、**be**、**cs6**、**cs7** 或者 **ef**。有关这些本地优先级名称所对应的具体含义参见第 10 章 10.1.3 节 （2）*queue-index*：表示队列的索引，S2700SI /2700EI/2710SI 系列交换机**仅支持 4 个队列**，索引号为 0～3 的整数 可使用 **undo qos local-precedence-queue-map** 命令恢复本地优先级和队列之间的映射关系为表 10-7 右边部分的缺省映射关系 【说明】内部优先级和入队列之间的映射关系仅在接口入方向上起作用，即映射关系影响报文流入队列操作。通过配置内部优先级和队列之间的映射关系，设备依据内部优先级和队列之间的映射关系将报文送入指定队列
（可选）管理优先级映射	7	**display qos local-precedence-queue-map** 例如：[HUAWEI] **display qos local-precedence-queue-map**	查看内部优先级到入队列的映射关系。在上一步配置了本地优先级和入队列之间的映射关系后，可通过本命令查看配置信息

【示例 】在S2700设备上配置了本地优先级和队列之间的映射关系后，查看相关配置信息。

<HUAWEI> system-view

[HUAWEI] qos local-precedence-queue-map af3 2

[HUAWEI] display qos local-precedence-queue-map

Current configurations of mapping between local-precedence and queue:

local-precedence value: be   queue index: 0
local-precedence value: af1   queue index: 0
local-precedence value: af2   queue index: 1
local-precedence value: af3   queue index: 2
local-precedence value: af4   queue index: 2
local-precedence value: ef   queue index: 2
local-precedence value: cs6   queue index: 3
local-precedence value: cs7   queue index: 3

11.1.2 其他 S2700/3700、S5700SI/5700EI/5700LI/5700S-LI 系列优先级映射配置与管理

在 S2700-52P-EI/2700-52P-PWR-EI/3700SI/3700EI，以及 S5700SI/5700EI/5700LI/5700S-LI系列交换机中所支持的优先级信任模式相同，即全面支持802.1p、IP和DSCP三种优先级，支持 “DSCP优先级到802.1p、DSCP、丢弃优先级的映射”和“IP优先级到802.1p、IP优先级的映射”两种优先级映射模式。

1．基本配置任务

在 S2700-52P-EI/2700-52P-PWR-EI/3700SI/3700EI，以及 S5700SI/5700EI/5700LI/5700S-LI系列交换机的优先级映射配置任务如下。

（1）配置端口信任的报文优先级。

（2）（可选）配置端口优先级（仅用于不带**VLAN**标签的报文优先级映射）。

（3）配置DSCP优先级到802.1p、新DSCP、DP（丢弃）优先级之间的映射。

（4）配置IP优先级到802.1p、新IP优先级之间的映射。

（5）（可选）配置内部优先级和队列之间的映射关系（仅当需要改变缺省的映射关系时配置）。

2．具体配置步骤

以上五项配置任务的具体配置步骤如表11-2所示。那些在接口视图下进行的配置，如果需要在多个端口下配置相同的配置，则可选择在对应端口组视图下进行配置，以减少配置工作量。

表11-2 S2700-52P-EI/2700-52P-PWR-EI/3700SI/3700EI/5700SI/5700EI/5700LI/5700S-LI系列优先级映射配置步骤

配置任务	步骤	命令	说明
配置端口信任的报文优先级	1	**system-view** 例如：<HUAWEI> **system-view**	进入系统视图
	2	**interface** *interface-type interface-number* 例如：[HUAWEI] **interface ethernet 0/0/1**	键入要配置端口信任的报文优先级的接口，进入接口视图
	3	**trust** { **8021p** | **dscp** | **ip-precedence** } 例如：[HUAWEI-Ethernet0/0/1] **trust 8021p**	配置端口的优先级信任模式。命令中的选项说明如下。 （1）**8021p**：多选一选项，指定端口信任 802.1p 优先级，使端口根据 802.1p 优先级对应的映射关系进行映射处理 （2）**dscp**：多选一选项，指定端口信任 DSCP 优先级，使端口根据 DSCP 优先级对应的映射关系进行映射处理

（续表）

配置任务	步骤	命令	说明
配置端口信任的报文优先级	3	**trust** { **8021p** | **dscp** | **ip-precedence** } 例如：[HUAWEI-Ethernet0/0/1] **trust 8021p**	（3）**ip-precedence**：多选一选项，指定端口信任IP优先级，使端口根据IP优先级对应的映射关系进行映射处理 【说明】当在同一接口下可同时配置 **trust dscp** 和 **trust 8021p** 命令，此时如果报文为IP报文，则接口信任报文的DSCP优先级；如果报文为非IP报文，则接口信任报文的802.1p优先级。但同一接口下不可同时配置 **trust dscp** 和 **trust ip-precedence** 命令 缺省情况下，不信任任何报文优先级，可用 **undo trust** 命令取消端口的信任优先级配置，即取消对报文按照某类优先级进行的映射，报文都进入队列0且报文的802.1p值被设置为0
（可选）配置端口优先级	4	**port priority** *priority-value* 例如：[HUAWEI-Ethernet0/0/1] **port priority** 1	配置端口的缺省802.1p优先级，取值范围为0～7的整数，取值越大优先级越高。如果当前接口已加入Eth-Trunk，本命令不可用，其他说明参见上节表11-1第4步
配置DSCP优先级与其他优先级的映射关系	5	**quit** 例如：[HUAWEI-Ethernet0/0/1] **quit**	退出接口视图，返回系统视图
	6	**qos map-table** { **dscp-dot1p** | **dscp-dp** | **dscp-dscp** } 例如：[HUAWEI] **qos map-table dscp-dot1p**	进入DSCP映射表视图，**仅支持DSCP到802.1p优先级的单向映射**。命令中的选项说明如下。 （1）**dscp-dot1p**：多选一选项，指定进入dscp-dot1p视图，即从DSCP优先级到802.1p优先级的映射视图，配置DSCP到802.1p优先级映射时选择，对接收的报文按照所携带的DSCP优先级重标记报文的8021.p优先级 （2）**dscp-dp**：多选一选项，指定进入dscp-dp视图，即从DSCP优先级到丢弃优先级的映射视图，配置DSCP到丢弃优先级映射时选择，对接收的报文按照所携带的DSCP优先级重标记报文的丢弃优先级 （3）**dscp-dscp**：多选一选项，指定进入dscp-dscp视图，即从DSCP优先级到DSCP优先级的映射视图，配置DSCP到DSCP优先级映射时选择，对接收的报文按照所携带的DSCP优先级重标记报文的DSCP优先级 **具体要进行何种映射，要视本表第3步所配置的端口优先级信任模式而定，信任哪种模式就可以把报文中携带的优先级映射成哪种的优先级**
配置DSCP优先级与其他优先级的映射关系	7	**input** { *input-value1* [**to** *input-value2*] &<1-10> } **output** *output-value* 例如：[HUAWEI-dscp-dot1p] **input** 0 **to** 15 **output** 0	配置DSCP表中的映射关系（先需要通过上一步进入到对应的映射表视图），可以修改DSCP表中DSCP到802.1p、DSCP到DP、DSCP到DSCP的映射关系。命令中的参数说明如下。 （1）*input-value1*：指定建立优先级映射表时输入的起始DSCP优先级值，取值范围为0～63的整数 （2）*input-value2*：可选参数，指定建立优先级映射表时输入的结束DSCP优先级值，取值范围也为0～63的整数，但要大于 *input-value1* 值。它和 *input-value1* 共同确定一个DSCP优先级值范围

（续表）

<table>
<tr><th>配置任务</th><th>步骤</th><th>命令</th><th>说明</th></tr>
<tr><td>配置 DSCP 优先级与其他优先级的映射关系</td><td>7</td><td>input { input-value1 [to input-value2] &<1-10> } output output-value
例如：[HUAWEI-dscp-dot1p] input 0 to 15 output 0</td><td>（3）output-value：指定输出的 802.1p 优先级、丢弃优先级或新的 DSCP 值。取值范围取决于当前映射表视图。
➢ 在 dscp-dot1p 视图下的取值范围为 0～7 的整数
➢ 在 dscp-dp 视图下的取值范围为 0～2 的整数：丢弃优先级 0 对应报文颜色 green；丢弃优先级 1 对应报文颜色 yellow；丢弃优先级 2 对应报文颜色 red
➢ 在 dscp-dscp 视图下的取值范围为 0～63 的整数
缺省情况下，DSCP 到 802.1p 的映射关系如表 11-3 所示，DSCP 到 DP 的映射关系如表 11-4 所示，DSCP 到 DSCP 的映射关系如表 11-5 所示，可用 undo input { all | input-value1 [to input-value2] &<1-10> }命令恢复缺省情况</td></tr>
<tr><td></td><td>8</td><td>quit
例如：[HUAWEI-dscp-dot1p] quit</td><td>退出 DSCP 映射表视图，返回系统视图</td></tr>
<tr><td rowspan="2">配置 IP 优先级与其他优先级的映射关系</td><td>9</td><td>qos map-table { ip-pre-dot1p | ip-pre-ip-pre }
例如：[HUAWEI] qos map-table ip-pre-dot1p</td><td>进入 IP 优先级映射表视图。命令中的选项说明如下。
（1）ip-pre-dot1p：二选一选项，指定进入 ip-pre-dot1p 视图，即从 IP 优先级到 802.1p 优先级的映射视图，配置 IP 优先级到 802.1p 优先级映射时选择，对接收的报文按照所携带的 IP 优先级重标记报文的 8021.p 优先级
（2）ip-pre-ip-pre：二选一选项，指定进入 ip-pre-ip-pre 视图，即从 IP 优先级到 IP 优先级的映射视图，在配置 IP 优先级到新 IP 优先级映射时选择，对接收的报文按照所携带的 IP 优先级重标记报文的 IP 优先级
具体要进行何种映射，要视本表第 3 步所配置的端口优先级信任模式而定，信任哪种模式就可以把报文中携带的优先级映射成哪种的优先级</td></tr>
<tr><td>10</td><td>input input-value1 [to input-value2] output output-value
例如：[HUAWEI-ip-pre-dot1p] input 0 to 7 output 0</td><td>配置 IP 优先级表中的映射关系（先需要通过上一步进入到对应的映射表视图），可以修改 IP 优先级表中 IP 优先级到丢弃优先级，IP 优先级到 IP 优先级的映射关系。命令中的参数说明如下。
（1）input-value1：指定建立优先级映射表时输入的起始 IP 优先级值，取值范围为 0～7 的整数
（2）input-value2：可选参数，指定建立优先级映射表时输入的结束 IP 优先级值，取值范围为 0～7 的整数，但要大于 input-value1 值。它和 input-value1 共同确定一个 IP 优先级值范围
缺省情况下，IP 优先级到丢弃优先级和 IP 优先级的映射关系都是 0、1…7 依次对应的，可用 undo input { all | input-value1 [to input-value2] }命令恢复缺省情况</td></tr>
</table>

（续表）

配置任务	步骤	命令	说明
（可选）配置内部优先级和队列之间的映射关系	11	**quit** 例如：[HUAWEI-dscp-dot1p] **quit**	退出 IP 映射表视图，返回系统视图
	12	**qos local-precedence-queue-map** *local-precedence queue-index* 例如：[HUAWEI] **qos local-precedence-queue-map** af3 2	配置内部优先级和队列之间的映射关系。其他说明参见上节表 11-1 中的第 6 步
（可选）管理优先级映射	13	**display qos map-table** [**dscp-dot1p** | **dscp-dp** | **dscp-dscp** | **ip-pre-dot1p** | **ip-pre-ip-pre**] 例如：[HUAWEI] **display qos map-table dscp-dp**	查看当前的各种优先级间的映射关系。命令中的选项可参见本表中前面 **qos map-table** 命令中的对应选项说明。如果没有指定任何可选项，将显示 DSCP 到 Dot1p、DSCP 到 DP、DSCP 到 DSCP、IP 优先级到 Dot1p、IP 优先级到 IP 优先级的全部映射关系
	14	**display qos local-precedence-queue-map** 例如：[HUAWEI]**display qos local-precedence-queue-map**	查看内部优先级到队列的映射关系。在前面配置了本地优先级和队列之间的映射关系后，可通过本命令查看配置信息

表11-3 缺省情况下的DSCP到802.1p映射关系表

Input DSCP	802.1p	Input DSCP	802.1p	Input DSCP	802.1p	Input DSCP	802.1p
0	0	16	2	32	4	48	6
1	0	17	2	33	4	49	6
2	0	18	2	34	4	50	6
3	0	19	2	35	4	51	6
4	0	20	2	36	4	52	6
5	0	21	2	37	4	53	6
6	0	22	2	38	4	54	6
7	0	23	2	39	4	55	6
8	1	24	3	40	5	56	7
9	1	25	3	41	5	57	7
10	1	26	3	42	5	58	7
11	1	27	3	43	5	59	7
12	1	28	3	44	5	60	7
13	1	29	3	45	5	61	7
14	1	30	3	46	5	62	7
15	1	31	3	47	5	63	7

表11-4 缺省情况下的DSCP到DP映射关系表

Input DSCP	DP	Input DSCP	DP	Input DSCP	DP	Input DSCP	DP
0	0	16	0	32	0	48	0
1	0	17	0	33	0	49	0
2	0	18	0	34	0	50	0
3	0	19	0	35	0	51	0

（续表）

Input DSCP	DP	Input DSCP	DP	Input DSCP	DP	Input DSCP	DP
4	0	20	0	36	0	52	0
5	0	21	0	37	0	53	0
6	0	22	0	38	0	54	0
7	0	23	0	39	0	55	0
8	0	24	0	40	0	56	0
9	0	25	0	41	0	57	0
10	0	26	0	42	0	58	0
11	0	27	0	43	0	59	0
12	0	28	0	44	0	60	0
13	0	29	0	45	0	61	0
14	0	30	0	46	0	62	0
15	0	31	0	47	0	63	0

表11-5 缺省情况下的DSCP到DSCP映射关系表

Input DSCP	DSCP	Input DSCP	DSCP	Input DSCP	DSCP	Input DSCP	DSCP
0	0	16	16	32	32	48	48
1	1	17	17	33	33	49	49
2	2	18	18	34	34	50	50
3	3	19	19	35	35	51	51
4	4	20	20	36	36	52	52
5	5	21	21	37	37	53	53
6	6	22	22	38	38	54	54
7	7	23	23	39	39	55	55
8	8	24	24	40	40	56	56
9	9	25	25	41	41	57	57
10	10	26	26	42	42	58	58
11	11	27	27	43	43	59	59
12	12	28	28	44	44	60	60
13	13	29	29	45	45	61	61
14	14	30	30	46	46	62	62
15	15	31	31	47	47	63	63

【示例 1】配置 DSCP表中的映射关系，将 DSCP的 0到 7级映射成 802.1p的0级。

<HUAWEI> system-view

[HUAWEI] qos map-table dscp-dot1p

[HUAWEI-dscp-dot1p] input 0 to 7 output 0

【示例 2】配置IP优先级表中的映射关系：IP优先级的0到3级映射成Dotq1的0级。

<HUAWEI> system-view

[HUAWEI] qos map-table ip-pre-dot1p

[HUAWEI-ip-pre-dot1p] input 0 to 3 output 0

11.1.3 优先级映射配置示例（一）

本示例适用于 S2700-52P-EI/2700-52P-PWR-EI/3700SI/3700EI/5700SI/5700EI/5700LI/ 5700S-LI系列交换机。本示例拓扑结构如图11-1所示，SwitchA和SwitchB都与路由器互连，企业分支机构 1 和企业分支机构 2 可经由 LSW1 和 LSW2 访问核心网络（Core Network）。现由于企业分支机构1需要得到更好的QoS保证，因此可将来自企业分支机构1的数据报文DSCP优先级映射为45，将来自企业分支机构2的数据报文DSCP优先级映射为30。当拥塞发生时，Router优先处理DSCP优先级高的报文。

1．基本配置思路分析

本示例总体上的配置思路如下。

（1）按照图示要求在各交换机上创建所需的VLAN，配置各接口为对应的类型，并加入对应的VLAN中，使企业都能够访问网络。

（2）在SwitchA和SwitchB上配置DSCP优先级信任，以及DSCP到DSCP的优先级映射关系，将来自企业分支机构 1 的数据报文优先级映射为 45，将来自企业分支机构 2 的数据报文优先级映射为 30，以实现为来自两个分支机构的报文提供差分化服务。

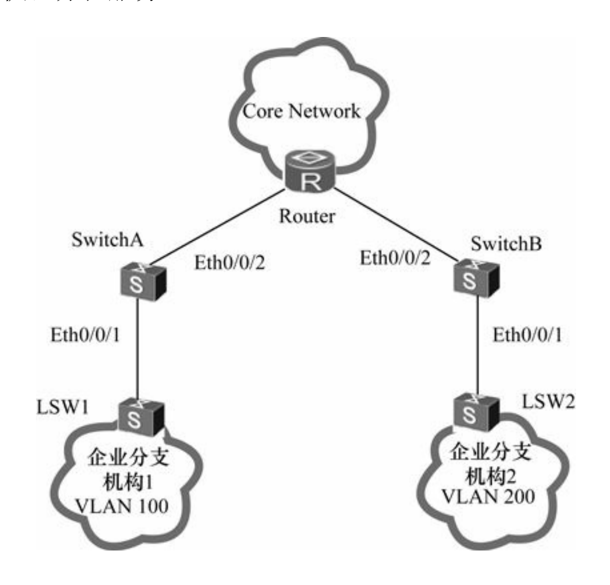


图11-1 优先级映射配置示例一拓扑结构

2．具体配置步骤

下面仅介绍SwitchA和SwitchB上的配置。

SwitchA上的配置：

（1）创建VLAN100。

```
<HUAWEI>system-view
[HUAWEI] sysname SwitchA
[SwitchA] vlan 100
```

（2）配置Eth0/0/1、Eth0/0/2端口类型均为trunk，并加入VLAN100。

```
[SwitchA] interface ethernet 0/0/1
[SwitchA-Ethernet0/0/1] port link-type trunk
[SwitchA-Ethernet0/0/1] port trunk allow-pass vlan 100
[SwitchA-Ethernet0/0/1] quit
[SwitchA] interface ethernet 0/0/2
[SwitchA-Ethernet0/0/2] port link-type trunk
[SwitchA-Ethernet0/0/2] port trunk allow-pass vlan 100
[SwitchA-Ethernet0/0/2] quit
```

（3）配置Eth0/0/1、Eth0/0/2端口信任报文的DSCP优先级。

```
[SwitchA] interface ethernet 0/0/1
[SwitchA-Ethernet0/0/1] trust dscp
[SwitchA-Ethernet0/0/1] quit
[SwitchA] interface ethernet 0/0/2
[SwitchA-Ethernet0/0/2] trust dscp
```

[SwitchA-Ethernet0/0/2] quit

（4）配置 DSCP到 DSCP的优先级映射，把报文中的 DSCP优先级全部重标记为45。

[SwitchA] qos map-table dscp-dscp

[SwitchA-dscp-dscp] input 0 to 63 output 45

SwitchB上的配置：

（1）创建VLAN200。

<HUAWEI>system-view

[HUAWEI] sysname SwitchB

[SwitchB] vlan 200

（2）配置Eth0/0/1、Eth0/0/2端口类型为trunk，并加入VLAN200。

[SwitchB] interface ethernet 0/0/1

[SwitchB-Ethernet0/0/1] port link-type trunk

[SwitchB-Ethernet0/0/1] port trunk allow-pass vlan 200

[SwitchB-Ethernet0/0/1] quit

[SwitchB] interface ethernet 0/0/2

[SwitchB-Ethernet0/0/2] port link-type trunk

[SwitchB-Ethernet0/0/2] port trunk allow-pass vlan 200

[SwitchB-Ethernet0/0/2] quit

（3）配置Eth0/0/1、Eth0/0/2端口信任报文的DSCP优先级。

[SwitchB] interface ethernet 0/0/1

[SwitchB-Ethernet0/0/1] trust dscp

[SwitchB-Ethernet0/0/1] quit

[SwitchB] interface ethernet 0/0/2

[SwitchB-Ethernet0/0/2] trust dscp

[SwitchB-Ethernet0/0/2] quit

（4）配置DSCP到DSCP的优先级映射，把报文中的DSCP优先级全部重标记为30，优先级次于VLAN100中的报文。

[SwitchB] qos map-table dscp-dscp

[SwitchB-dscp-dscp] input 0 to 63 output 30

配置好后，可在两交换机上使用对应的**display qos map-table** 命令查看优先级映射信息，也可在对应的接口视图下通过 **display this** 命令查看接口上的优先级映射配置信息，以验证配置结果。

11.1.4 S5700HI/5710EI/6700/7700/9300/9300E/9700 系列优先级映射配置与管理

在S5700HI/5710EI/6700/7700/9300/9300E/9700系列交换机中，仅支持“信任报文的802.1p优先级”和“信任报文的DSCP优先级”这两种优先级信任模式。但这些交换机 支持以下多种优先级映射模式（此处不考虑EXP优先级）。

（1）在接口入方向上的802.1p优先级到PHB行为/颜色映射。

（2）在接口出方向上的PHB行为/颜色到802.1p优先级映射。

（3）在接口入方向上的DSCP优先级到PHB行为/颜色。

（4）在接口出方向上的PHB行为/颜色到DSCP优先级。

这里有一个难点就是这些交换机开始支持DiffServ域。DiffServ域其实与其他域类似，如在本书第8章中所介绍的MST域，都是用来指定一种配置的生效范围。DiffServ域用于定义报文的优先级（支持802.1p优先级和DSCP优先级）和PHB行为之间的映射关系，并为报文标记颜色，用于进行拥塞管理和拥塞避免。这样一来，就可以在整个网络中配置不同的优先级和PHB行为之间的映射关系，作用于不同DiffServ域内的报文。可以在交换机的端口上绑定对应DiffServ域中配置的优先级和PHB行为映射关系。

1．基本配置任务

总体来说，S5700HI/5710EI/6700/7700/9300/9300E/9700系列交换机的优先级映射配置任务包括以下几个方面。

（1）配置端口信任的报文优先级。

（2）（可选）配置端口优先级。

（3）创建DiffServ域并配置优先级映射关系。

当设备作为DiffServ域和其他网络的边界节点时，需要配置内部优先级和外部优先级的相互映射关系。当业务流流入设备时将报文携带的优先级信息映射到相应的PHB行为/颜色，并在设备内部根据报文的 PHB 行为进行拥塞管理，根据报文的颜色进行拥塞避免；当业务流流出设备时将报文的 PHB 行为/颜色映射为相应的优先级，对端设备根据报文的优先级提供相应的QoS服务。

（4）应用DiffServ域。

当需要根据 DiffServ 域中定义的映射关系，对出/入设备的报文进行优先级到 PHB行为/颜色之间的映射操作时，可以将 DiffServ 域绑定到报文的出/入接口，系统会根据DiffServ域中的映射关系进行报文优先级与PHB行为/颜色之间的映射。

2．具体配置步骤

以上4项配置任务的具体配置步骤如表11-6所示。对于那些接口视图下的配置，如果需要在多个端口下配置相同的配置，则可选择在对应的端口组视图下进行配置，以减轻配置工作量。

表11-6 S5700HI/5710EI/6700/7700/9300/9300E/9700系列优先级映射配置步骤

配置任务	步骤	命令	说明
配置端口信任的报文优先级	1	**system-view** 例如：<HUAWEI> **system-view**	进入系统视图
	2	**interface** *interface-type interface-number* 例如：[HUAWEI] **interface ethernet 0/0/1**	键入要配置端口信任的报文优先级的接口，进入接口视图

（续表）

配置任务	步骤	命令	说明
配置端口信任的报文优先级	3	**trust** { **8021p** { **inner** | **outer** } | **dscp** } 例如：[HUAWEI-Ethernet0/0/1] **trust 8021p inner**	配置端口的优先级信任模式。命令中的选项说明如下： （1）**8021p** { **inner** | **outer** }：二选一选项，指定端口信任 802.1p 优先级，如果选择了二选一选项 **inner**，则指定对报文按照内层 802.1p 优先级进行映射；如果选择了二选一选项 **outer**，则指定对报文按照外层 802.1p 优先级进行映射 （2）**dscp**：二选一选项，指定端口信任 DSCP 优先级，使端口根据 DSCP 优先级对应的映射关系进行映射处理 （3）**ip-precedence**：多选一选项，指定端口信任 IP 优先级，使端口对报文按照 DSCP 优先级进行映射 本命令为覆盖式命令，即**在同一接口视图下多次执行该命令配置后，按最后一次配置生效** 缺省情况下，根据外层 802.1p 优先级对应的映射关系进行映射处理，可用 **undo trust** 命令取消对报文按照某类优先级进行的映射
（可选）配置端口优先级	4	**port priority** *priority-value* 例如：[HUAWEI-Ethernet1/0/1] **port priority** 1	配置端口的缺省 802.1p 优先级（**但 S6700 系列不支持配置端口优先级**），取值范围为 0～7 的整数，值越大优先级越高。如果当前接口已加入 Eth-Trunk，本命令不可用，其他说明参见 11.1.1 节表 11-1 第 4 步 **【注意】**S6700 系列，S7700 系列的 ES1D2X40SFC0 单板、ES1D2L02QFC0 单板，S9300 系列的 LE0DX40SFC00 单板、LE1D2L02QFC0 单板，S9300E 系列的 LE0DX40SFC00 单板、LH2D2L02QFC0 单板，以及 S9700 系列的 EH1D2X40SFC0 单板、EH1D2L02QFC0 单板均不支持配置端口优先级
创建 DiffServ 域并配置优先级映射关系	5	**quit** 例如：[HUAWEI-Ethernet0/0/1] **quit**	退出接口视图，返回系统视图
	6	**diffserv domain** { **default** | *ds-domain-name* } 例如：[HUAWEI] **diffserv domain** ds1	创建 DiffServ 域并进入 DiffServ 域视图。命令中的参数和选项说明如下。 （1）**default**：二选一选项，指定进入系统预先设定的缺省 DiffServ 域视图 （2）*ds-domain-name*：二选一参数，指定新创建的 DiffServ 域名称，为 1～31 个字符，不支持空格，不区分大小写，且不能为“n”、“no”、“non”、“none” 设备中缺省存在一个名为 **default** 的 DiffeServ 域，除了这个域设备最多允许创建 7 个 DiffServ 域。对于预先设定的 **default** 域，用户只能修改其映射关系，不能删除，但可用 **undo diffserv domain** *ds-domain-name* 删除新建的指定 DiffServ 域

（续表）

配置任务	步骤	命令	说明	
创建DiffServ域并配置优先级映射关系	7	**8021p-inbound** *8021p-value* **phb** *service-class* [**green** | **yellow** |**red**] 例如：[HUAWEI-dsdomain-ds1] **8021p-inbound** 2 **phb** af1 **yellow**	配置“802.1p 优先级到 PHB 行为/颜色”优先级映射模式，在接口入方向，将 VLAN 报文的 802.1p 优先级映射为 PHB 行为，并为报文着色。命令中的参数和选项说明如下（颜色仅用在流量控制时识别是否丢包，对内部优先级与队列的映射关系没有影响）。 （1）*8021p-value*：指定要建立映射关系的 VLAN 报文的 802.1p 优先级值，取值范围为 0～7 的整数，值越大优先级越高 （2）*service-class*：指定所映射的 PHB 行为，取值可以为 BE、AF1～AF4、EF、CS6 或 CS7（优先级依次增高），并与设备的端口队列 0～7 依次对应 （3）**green**：多选一可选项，指定将报文标记为绿色 （4）**yellow**：多选一可选项，指定将报文标记为黄色 （5）**red**：多选一可选项，指定将报文标记为红色 将 DiffServ 域绑定到报文的入接口后，根据报文的 PHB 行为将其送入对应的端口队列，进行拥塞管理，配置丢弃模板以后根据报文的颜色进行拥塞避免 缺省情况下，DiffServ 域中接口入方向上 VLAN 报文的 802.1p 优先级和 PHB 行为/颜色之间的映射关系**如表 11-7 所示**，可用 **undo 8021p-inbound** [*8021p-value*] 命令恢复缺省的映射关系，如不指定可选参数 *8021p-value*，将取消所有 802.1p 值与服务等级的对应关系的配置	根据需要选择执行一个或多个配置
		8021p-outbound *service-class* { **green** | **yellow** | **red** } **map** *8021p-value* 例如：[HUAWEI-dsdomain-ds1] **8021p-outbound** af1 **yellow map** 2	配置“PHB 行为/颜色到 802.1p 优先级”优先级映射模式，在接口出方向，将 PHB 行为/颜色映射为 VLAN 报文的 802.1p 优先级。命令中的参数和选项说明参见上一步的 **8021p-inbound** 命令对应的参数和选项说明 将 DiffServ 域绑定到报文的入接口后，根据报文的 PHB 行为将其送入对应的端口队列，进行拥塞管理，配置丢弃模板以后根据报文的颜色进行拥塞避免 缺省情况下，DiffServ 域中接口出方向上 VLAN 报文的 PHB 行为/颜色和 802.1p 优先级之间的映射关系**如表 11-8 所示**，可用 **undo 8021p-outbound** [*service-class* { **green** | **yellow** | **red** }] 命令恢复缺省的映射关系	
		ip-dscp-inbound *dscp-value* **phb** *service-class* [**green** | **yellow** |**red**] 例如：[HUAWEI-dsdomain-ds1] **ip-dscp-inbound** 8 **phb** af1 **yellow**	配置“DSCP 优先级到 PHB 行为/颜色”优先级映射模式，在接口入方向将 IP 报文的 DSCP 优先级映射为 PHB 行为，并为报文着色。命令中的参数和选项说明如下。 （1）*dscp-value*：指定要建立映射关系的 IP 报文的 DSCP 优先级值，取值范围为 0～63 的整数，值越大优先级越高 （2）*service-class*：指定所映射的 PHB 行为，取值可以为 BE、AF1～AF4、EF、CS6 或 CS7（优先级依次增高），并与设备的端口队列 0～7 依次对应	

（续表）

配置任务	步骤	命令	说明	
创建 DiffServ 域并配置优先级映射关系	7	**ip-dscp-inbound** *dscp-value* **phb** *service-class* [**green** | **yellow** |**red**] 例如：[HUAWEI-dsdomain-ds1] **ip-dscp-inbound** 8 **phb** af1 **yellow**	（3）**green**：多选一可选项，指定将报文标记为绿色 （4）**yellow**：多选一可选项，指定将报文标记为黄色 （5）**red**：多选一可选项，指定将报文标记为红色 将 DiffServ 域绑定到报文的入接口后，根据报文的 PHB 行为将其送入对应的端口队列，进行拥塞管理，配置丢弃模板以后根据报文的颜色进行拥塞避免 缺省情况下，DiffServ 域中接口入方向上 IP 报文的 DSCP 优先级和 PHB 行为/颜色之间的映射关系**如表 11-9 所示**，可用 **undo ip-dscp-inbound** [*dscp-value*] 命令恢复缺省的映射关系，如果不指定可选参数 *dscp-value*，将取消所有 dscp 值与服务等级映射关系的配置	根据需要选择执行一个或多个配置
		ip-dscp-outbound *service-class* { **green** | **yellow** | **red** } **map***dscp-value* 例如：[HUAWEI-dsdomain-ds1] **ip-dscp-outbound** af1 **yellow map** 8	配置"PHB 行为/颜色到 DSCP 优先级"优先级映射模式，在接口出方向，将 PHB 行为/颜色映射为 IP 报文的 DSCP 优先级。命令中的参数和选项说明参见上一步的 **ip-dscp-inbound** 命令对应的参数和选项说明 将 DiffServ 域绑定到报文的出接口后，下游设备将根据报文的 DSCP 优先级进行调度 缺省情况下，DiffServ 域中接口出方向上 IP 报文的 PHB 行为/颜色和 DSCP 优先级之间的映射关系**如表 11-10 所示**，可用 **undo ip-dscp-outbound** [*service-class* { **green** | **yellow** | **red** }]命令恢复缺省的映射关系	
应用 DiffServ 域	8	**quit** 例如：[HUAWEI-dscp-dot1p] **quit**	退出 DSCP 映射表视图，返回系统视图	
	9	**interface** *interface-type interface-number* 例如：[HUAWEI] **interface** gigabitethernet 0/0/1	键入要绑定 DiffServ 域的交换机端口，进入接口视图	
	10	**trust upstream** { *ds-domain-name* | **default** | **none** } 例如：[HUAWEI-GigabitEthernet0/0/1] **trust upstream** ds1	在以上接口上绑定指定的 DiffServ 域。绑定后，系统会根据 DiffServ 域中的映射关系将流经该接口的报文优先级与 PHB 行为或者报文颜色进行映射。命令中的参数和选项说明如下。 （1）*ds-domain-name*：多选一参数，指定要绑定的 DiffServ 域的域名 （2）**default**：多选一选项，指定绑定缺省的 DiffServ 域 （3）**none**：多选一选项，指定不信任报文优先级，取消接口上的 PHB 映射功能，配置后，系统对出入接口的报文都不进行 PHB 映射	

（续表）

配置任务	步骤	命令	说明
应用DiffServ域	10	**trust upstream** { *ds-domain-name* | **default** | **none** } 例如：[HUAWEI-GigabitEthernet0/0/1] **trust upstream** ds1	【说明】当需要根据DiffServ域中定义的映射关系，对来自上游设备的报文进行优先级到PHB行为之间的映射操作时，可以通过本命令将DiffServ域应用到报文的入接口；当需要根据DiffServ域中定义的映射关系，对流向下游设备的报文进行PHB行为到优先级之间的映射操作时，可以通过本命令将DiffServ域应用到报文的出接口。但本命令为覆盖式命令，即在同一接口视图下多次执行该命令配置后，**按最后一次配置生效** 缺省情况下，接口上不应用任何DiffServ域，可用**undo trust upstream**命令恢复缺省配置。如果要修改接口下绑定的DiffServ域，必须先执行**undo trust upstream**命令删除已绑定的DiffServ域，再执行本命令重新应用新的DiffServ域
	11	**undo qos phb marking enable** 例如：[HUAWEI-GigabitEthernet0/0/1] **undo qos phb marking enable**	（可选）取消对接口出方向的报文进行PHB映射（不影响系统对入接口的报文进行PHB映射）。通常在作为DS域边界节点的设备上配置本命令关闭与非DS域设备连接的接口的PHB映射功能。 缺省情况下，对接口出方向的报文进行PHB映射
（可选）管理优先级映射	12	**quit** 例如：[HUAWEI-dscp-dot1p] **quit**	退出IP映射表视图，返回系统视图
	13	**display diffserv domain** [**all** | **name** *ds-domain-name*] 例如：[HUAWEI] **display diffserv domain name** d1	查看DiffServ域的配置信息。当用户创建了DiffServ域并对其中的映射关系进行配置后，可以通过本命令查看该DiffServ域的配置信息。如果不指定所有可选参数，该命令将显示设备上已创建的所有DiffServ域的概要信息
	14	**display qos local-precedence-queue-map** 例如：[HUAWEI] **display qos local-precedence-queue-map**	查看内部优先级到队列的映射关系，**但S6700不支持本命令**。在前面步骤配置了本地优先级和队列之间的映射关系后，可通过本命令查看配置信息

注意

在S7700/9300/9700系列交换机中配置端口信任的报文优先级时应注意的地方。

（1）在S系列单板的报文入接口上配置trust 8021p inner命令时，实际使用的仍然是外层VLAN标签的802.1p优先级进行映射。

（2）在S系列单板的报文入接口上配置trust 8021p outer命令时，使用外层VLAN标签的802.1p优先级进行映射。如果报文不带VLAN标签，则按照端口缺省的802.1p优先级进入队列。

（3）在E系列单板和F系列单板的报文入接口中，系统按照实际配置对报文进行优先级映射。

（4）E系列单板和F系列单板的报文出接口上配置trust 8021p inner命令时，如果从接口发出的报文是双层 VLAN 标签，则根据 PHB/颜色映射到的外层 VLAN 标签的802.1p优先级会被写入外层VLAN标签的802.1p优先级字段，而不会写入内层VLAN标签的802.1p优先级字段。

（5）WAN接口板仅支持 trust dscp配置，且缺省不信任任何报文优先级。

表11-7 DiffServ域中接口入方向上802.1p优先级与PHB行为/颜色之间的缺省映射关系

802.1p 优先级	PHB 行为	颜色
0	BE	绿色
1	AF1	绿色
2	AF2	绿色
3	AF3	绿色
4	AF4	绿色
5	EF	绿色
6	CS6	绿色
7	CS7	绿色

从表11-7可以看出，缺省情况下，在接口入方向上，VLAN报文的“802.1p优先级与PHB行为的映射关系”与“802.1p优先级与内部优先级”的映射关系是完全一样的，并将所有报文颜色都标识为绿色。

表11-8 DiffServ域中接口出方向上PHB行为/颜色与802.1p优先级之间的缺省映射关系

PHB 行为	颜色	802.1p 优先级
BE	绿色	0
BE	黄色	0
BE	红色	0
AF1	绿色	1
AF1	黄色	1
AF1	红色	1
AF2	绿色	2
AF2	黄色	2
AF2	红色	2
AF3	绿色	3
AF3	黄色	3
AF3	红色	3
AF4	绿色	4
AF4	黄色	4
AF4	红色	4
EF	绿色	5
EF	黄色	5
EF	红色	5
CS6	绿色	6
CS6	黄色	6
CS6	红色	6
CS7	绿色	7
CS7	黄色	7
CS7	红色	7

从表11-8可以看出，缺省情况下，在接口出方向 上VLAN报文的“PHB行为/颜色与802.1p优先级的映射关系”中，“PHB行为与802.1p优先级的映射关系”仍是与“内部优先级与802.1p优先级的映射关系”一样，但每种802.1p优先级和PHB行为对应红、黄、绿三种颜色。

表11-9 DiffServ域中接口入方向上DSCP优先级和PHB行为/颜色之间的缺省映射关系

DSCP	PHB 行为	颜色	DSCP	PHB 行为	颜色
0	BE	绿色	32	AF4	绿色
1	BE	绿色	33	BE	绿色
2	BE	绿色	34	AF4	绿色
3	BE	绿色	35	BE	绿色
4	BE	绿色	36	AF4	黄色
5	BE	绿色	37	BE	绿色
6	BE	绿色	38	AF4	红色
7	BE	绿色	39	BE	绿色
8	AF1	绿色	40	EF	绿色
9	BE	绿色	41	BE	绿色
10	AF1	绿色	42	BE	绿色
11	BE	绿色	43	BE	绿色
12	AF1	黄色	44	BE	绿色
13	BE	绿色	45	BE	绿色
14	AF1	红色	46	BE	绿色
15	BE	绿色	47	BE	绿色
16	AF2	绿色	48	CS6	绿色
17	BE	绿色	49	BE	绿色
18	AF2	绿色	50	BE	绿色
19	BE	绿色	51	BE	绿色
20	AF2	黄色	52	BE	绿色
21	BE	绿色	53	BE	绿色
22	AF2	红色	54	BE	绿色
23	BE	绿色	55	BE	绿色
24	AF3	绿色	56	CS7	绿色
25	BE	绿色	57	BE	绿色
26	AF3	绿色	58	BE	绿色
27	BE	绿色	59	BE	绿色
28	AF3	黄色	60	BE	绿色
29	BE	绿色	61	BE	绿色
30	AF3	红色	62	BE	绿色
31	BE	绿色	63	BE	绿色

表11-10 DiffServ域中接口出方向上PHB行为/颜色和DSCP优先级之间的缺省映射关系

PHB 行为	颜色	DSCP
BE	绿色	0
BE	黄色	0
BE	红色	0
AF1	绿色	10
AF1	黄色	12
AF1	红色	14
AF2	绿色	18
AF2	黄色	20
AF2	红色	22
AF3	绿色	26
AF3	黄色	28
AF3	红色	30
AF4	绿色	34
AF4	黄色	36
AF4	红色	38
EF	绿色	46
EF	黄色	46
EF	红色	46
CS6	绿色	48
CS6	黄色	48
CS6	红色	48
CS7	绿色	56
CS7	黄色	56
CS7	红色	56

11.1.5 优先级映射配置示例（二）

本示例适用于S5700HI/5710EI/6700/7700/9300/9300E/9700系列交换机，拓扑结构如图11-2所示。Switch通过GE0/0/3端口与路由器互连，企业分支机构1和企业分支机构2可经由Switch和路由器访问网络。企业分支机构 1和企业分支机构 2的VLAN ID分别为100、200，且它们的报文802.1p值均为0。

现企业分支机构1需要更高的服务等级，以便得到更好的QoS保证。这时可通过定义DiffServ域，将来自企业分支机构 1的数据报文优先级映射为AF4 PHB行为，将来自企业分支机构 2的数据报文优先级映射为AF2 PHB行为，以提供差分服务（AF4的优先级要高于AF2的优先级）。

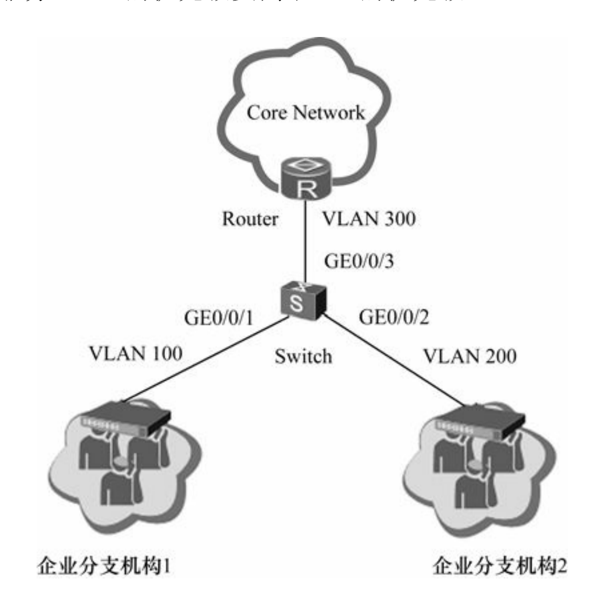


图11-2 优先级映射配置示例二拓扑结构

1．基本配置思路分析

本示例的配置也很简单，基本配置思路如下。

（1）在 Switch 交换机上创建所需 VLAN，并配置各接口类型，并加入对应的VLAN中。

（2）创建两个不同的DiffServ域，并分配将802.1p优先级值0映射为不同的PHB行为和颜色。其中一个映射为优先级更高的AF4 PHB行为，另一个映射为优先级较低的AF2 PHB行为。

（3）在Switch入GE0/0/1和GE0/0/2接口上对应绑定以上两个DiffServ域。

2．具体配置步骤

（1）批量创建所需要VLAN。

```
<HUAWEI> system-view
[HUAWEI] sysname Switch
[Switch] vlan batch 100 200 300
```

（2）将 GE0/0/1、GE0/0/2、GE0/0/3 端口类型均配置为 trunk，然后加入对应的VLAN中。

```
[Switch] interfacegigabitethernet0/0/1
[Switch-GigabitEthernet0/0/1] port link-type trunk
[Switch-GigabitEthernet0/0/1] port trunk allow-pass vlan 100
[Switch-GigabitEthernet0/0/1] quit
[Switch] interface gigabitethernet0/0/2
[Switch-GigabitEthernet0/0/2] port link-type trunk
[Switch-GigabitEthernet0/0/2] port trunk allow-pass vlan 200
[Switch-GigabitEthernet0/0/2] quit
[Switch] interface gigabitethernet0/0/3
[Switch-GigabitEthernet0/0/3] port link-type trunk
[Switch-GigabitEthernet0/0/3] port trunk allow-pass vlan 100 200 300
[Switch-GigabitEthernet0/0/3] quit
```

（3）在Switch上创建DiffServ域ds1、ds2，并配置将来自企业分支机构1和企业分支机构2报文中的802.1p优先级值0映射到不同的服务等级（分别为AF4和AF2）。

```
[Switch] diffserv domain ds1
[Switch-dsdomain-ds1] 8021p-inbound 0 phb af4 green
[Switch-dsdomain-ds1] quit
[Switch] diffserv domain ds2
[Switch-dsdomain-ds2] 8021p-inbound 0 phb af2 green
[Switch-dsdomain-ds2] quit
```

（4）将DiffServ域ds1和ds2分别绑定到接口GE0/0/1、GE0/0/2，使以上两种优先级映射在具体端口上应用。

```
[Switch] interface gigabitethernet0/0/1
[Switch-GigabitEthernet0/0/1] trust upstream ds1
[Switch-GigabitEthernet0/0/1] quit
[Switch] interface gigabitethernet0/0/2
```

```
[Switch-GigabitEthernet0/0/2] trust upstream ds2
[Switch-GigabitEthernet0/0/2] quit
```

11.2 流量监管和流量整形配置

流量监管和流量整形是通过监控进入网络的流量速率来限制流量及其资源的使用，保证更好地为用户提供服务。

11.2.1 流量监管配置综述

流量监管（TP）就是对流量进行控制，通过监督进入端口的流量速率，对超出部分的流量进行丢弃，使进入的流量被限制在一个合理的范围之内，从而保护网络资源和用户的利益。配置基于接口的流量监管后，设备将对流入接口上所有的业务流量进行流量监管；配置基于流的流量监管后，设备将对符合流分类规则的报文进行流量监管。具体流量监管原理参见第10章10.3.4节。

在所有华为S系列交换机中均具备以下几种流量监管功能（可同时配置 ）。

1．基于接口的流量监管

基于接口的流量监管是指对进入该接口的所有流量进行控制，而不区分具体报文的类型。监管的结果是丢弃超出速率限制的部分，使进入设备的该类流量被限制在一个合理的范围之内，从而保护网络资源。

2．基于管理网口的流量监管

当设备的管理网口由于恶意攻击、网络异常等原因导致流量过大时，会导致CPU占用率过高，进而影响系统正常运行。为了使系统正常运行，也需要对管理网口的流量进行限制。

S2700/3700系列不支持基于管理网口的流量监管。

3．基于流的流量监管

基于流的流量监管是指在设备上经过流分类后，对符合流分类的流量进行速率限制。通过监督进入设备的该类流量速率，丢弃超出速率限制的部分，使进入设备的该类流量被限制在一个合理的范围之内。基于流的流量监管采用双令牌桶技术，可实现出/入两个方向的端口限速和流限速。

S2700SI 系列交换机不支持基于流的流量监管 。S5700HI/5710EI/7700/9300/ 9300E/9700系列交换机支持对上行流作两次CAR，即在对符合流分类的上行流作完一次CAR 后，将所有的上行流聚合在一起再作一次共享 CAR，这里所有的上行流是指满足同一流策略中绑定了配置共享CAR的流行为的流分类的入方向业务流。

从配置方法上来看， S2700/3700/5700/6700 系列交换机的配置方法一样，而S7700/9300/9300E/9700系列交换机的配置方法一样。下面分别予以介绍。

11.2.2 配置流量监管

华为S系列交换机支持基于接口、基于管理网口和基于流这三种流量监管功能，用户可以选择配置所需的一种或全部的流量监管功能。下面分别介绍这三种流量监管功能的配置方法。

1．配置基于接口的流量监管

基于接口的流量监管是在交换机端口上应用监管策略，控制进入该端口的所有流量的速率，使端口接收的流量被限制在一个合理的范围之内（丢弃超出速率限制的部分）。但**S2700SI**和**S2710SI**系列交换机不支持对接口入方向配置流量监管 。在基于接口的流量监管配置方面，S2700/3700/5700/6700系列与S7700/9300/9300E/9700系列的配置方法有所不同。

（1）S2700/3700/5700/6700系列交换机基于接口的流量监管配置

在 S2700/3700/5700/6700 系列交换机中，是在对应的接口视图下使用 qos lr inbound cir cir-value [cbs cbs-value] 命令配置接口入方向上的流量监管速率（出方向上的流量控制是通过本章后面将要介绍的流量整形来实现的）。通过执行本命令，如果接口在入方向上收到报文的速率大于配置的流量监管速率，那么该报文将被丢弃。如果多个接口需要配置相同的流量监管速率，可通过接口组进行配置，以减少重复配置工作 。命令中的参数说明如下。

① cir-value：指定接口承诺信息速率（Committed Information Rate），即保证能够通过的平均速率，单位为kbit/s，但不同的接口类型取值范围不同，如标准以太网接口的取值范围为 64～100 000的整数，千兆以太网接口的取值范围为 64～1 000 000的整数，万兆以太网接口的取值范围为 64～10 000 000 的整数，40GE接口的取值范围为 64～40 000 000。

② cbs-value：可选参数，指定承诺突发尺寸（Committed Burst Size），即瞬间能够通过的承诺突发流量。单位为 byte（字节），取值范围为 4 000～4 294 967 295的整数。若不指定该参数：单令牌桶时，本参数的缺省值为参数cir-value取值的188倍；双令牌桶时，本参数的缺省值为cir-value参数取值的125倍。

本命令为覆盖式命令，即在同一接口多次配置流量整形参数后，按最后一次配置生效 。缺省情况下，接口入方向上不进行流量监管，即监管速率缺省为接口的最大带宽，可用**undo qos lr inbound** 命令恢复对应接口或接口组入方向上的流量监管速率为缺省值。

（2）在S7700/9300/9300E/9700系列交换机基于接口的流量配置

在 S7700/9300/9300E/9700 系列中，基于接口的流量配置需要先在系统视图下配置好CAR模板，然后在具体接口下配置监管速率。具体步骤如下。

① 在系统视图下通过qos car car-name｛cir cir-value [cbs cbs-value [pbs pbs-value] |pir pir-value [cbs cbs-valuepbs pbs-value]] ｝命令创建并配置 CAR模板，其中的参数说明如下。

● car-name：指定QoS CAR模板名称，取值范围为 1～31个字符。

● cir-value：二选一参数，指定承诺信息速率，即保证能够通过的平均速率，单位为kbit/s，取值范围为 64～4 294 967 295的整数。

● cbs-value：可选参数，指定承诺突发尺寸，即瞬间能够通过的承诺突发流量。单位为 byte（字节），取值范围为 4 000～4 294 967 295的整数。若不指定该参数：单令牌桶时，本参数的缺省值为参数cir-value取值的188倍；双令牌桶时，本参数的缺省值为cir-value参数取值的125倍。

● pir-value：二选一参数，指定峰值信息速率（Peak Information Rate），即最大能够通过的速率。单位为 kbit/s，取值范围为 64～4 294 967 295的整数，但本参数值必须大于等于cir-value参数值。

● pbs-value：可选参数，指定峰值突发尺寸（Peak Burst Size），即瞬间能够通过的峰值突发流量，单位为 byte，取值范围为 10 000～4 294 967 295的整数。但本参数值必须大于等于cbs-value参数值，缺省为pir-value值的125倍。

说明

当流量监管速率取值超过接口最大速率时，相当于没有对接口实行流量监管。请 根据接口的实际速率配置cir-value和pir-value参数的值小于接口速率。

当cbs-value值小于当前部署业务中单个报文的字节数时，将导致这些报文被直接丢弃。为避免报文颜色识别出现问题，建议配置pbs-value参数值大于cbs-value参数值。

报文的颜色由qos car中的 cbs cbs-value、pbs pbs-value参数共同确定。

● 报文的突发尺寸<cbs-value时，报文被标记为绿色。

● cbs-value≤报文的突发尺寸<pbs-value时，报文被标记为黄色。

● 报文的突发尺寸≥pbs-value时，报文被标记为红色。

说明

本命令创建的QoS CAR模板可以应用到以下具体场景中。

● 在流行为视图下，执行car car-name share命令对指定类别的业务流量进行流量监管。

● 在接口视图下，执行qos car inbound car-name命令对接口入方向上的所有报文进行流量监管。

● 在 VLAN 视图下应用 QoS CAR 模板时，执行 storm suppression broadcast car-name [share] 命令可对VLAN的上行广播流量进行流量监管；执行storm suppression multicast car-name [share] 命令可对 VLAN 的上行组播流量进行流量监管；执行unicast-suppression car-name [share] 命令可对VLAN的上行未知单播流量进行流量监管。

缺省情况下，系统未创建QoS CAR模板，可用undo qos car car-name命令删除指定的QoS CAR模板。

② 然后在具体接口视图下配置qos car inbound car-name命令，在接口上应用前面创建的 CAR 模板，以对流入该接口的所有业务流量实施流量监管。如果多个接口需要配置相同的QoS CAR，可通过端口组进行配置，以减少重复配置工作。

缺省情况下，接口上不应用任何QoS CAR模板，可用undo qos car inbound命令删除在对应接口或者接口组入方向上应用的QoS CAR模板。

【示例 1】在S2700/3700/5700/6700系列交换机上限制Eth0/0/1接口入方向上接收数据的承诺信息速率为20 000kbit/s，承诺突发尺寸为 375 000字节。

```
<HUAWEI> system-view
[HUAWEI] interface ethernet 0/0/1
[HUAWEI-Ethernet0/0/1] qos lr inbound cir 20000 cbs 375000
[HUAWEI-Ethernet0/0/1] quit
```

【示例 2】在S7700/9300/9700系列交换机上，对GE1/0/1接口入方向报文应用名为qoscar1 的 QoS CAR模板进行流量监管，限制该接口上接收数据的承诺速率为10 000kbit/s，承诺突发尺寸为 10 240字节。

```
<HUAWEI>system-view
[HUAWEI] qos car qoscar1 cir 10000 cbs 10240
[HUAWEI] interfacegigabitethernet 1/0/1
[HUAWEI-GigabitEthernet1/0/1] qos car inbound qoscar1
[HUAWEI-GigabitEthernet1/0/1] quit
```

可通过display qos car { all |name car-name }命令查看CAR模板的配置信息。

2．配置管理网口的流量监管

基于管理网口的流量监管（S2700/3700系列交换机不支持）是专门针对管理端口进行的流量监管，以防一些对管理端口（在 S5700/6700 系列交换机中为 meth 0/0/1， S7700/9300/9700系列交换机中为 ethernet 0/0/0）的恶意攻击，使得设备CPU占用率过高，进而影响系统正常运行。通过在管理端口上配置流量监管，限制由管理网口进入设备的流量速率，以保证系统正常运行。

管理网口的流量监管的配置也很简单，就是在对应的管理端口视图下使用 qos lr pps packets命令配置对于管理网口的限速功能，即限制管理网口的包速率。参数用来指定管理端口上的包速率，即每秒通过的报文数，单位为 pps，取值范围是（1～10 240）的整数。

【示例 3】限制S5700系列交换机管理网口MEth0/0/1的包速率为 1 000pps。

```
<HUAWEI>system-view
[HUAWEI] interface meth 0/0/1
```

[HUAWEI -MEth0/0/1] qos lr pps 1000

【示例 4】限制S7700系列交换机管理网口Ethernet0/0/0的包速率为 1 000pps。

<HUAWEI> system-view

[HUAWEI] interface ethernet 0/0/0

[HUAWEI-Ethernet0/0/0] qos lr pps 1000

3．配置基于流的流量监管

若需要对接口入方向某类流量进行控制时，就不能采用前面介绍的基于接口的流量监管了，而是需要采用此处介绍的基于流的流量监管。但**S2700SI**系列交换机不支持基于流的流量监管 。

基于流的流量监管其实就是一个配置和应用QoS策略的过程，其中包括“定义流分类”、“配置流行为”、“配置流策略”和“应用流策略”四大步骤，基本说明如下（具体在本章后面介绍）。

（1）定义流分类。可根据报文中的二层信息、三层信息、ACL对不同报文进行分类。 （2）配置流行为。创建流行为，并对不同分类的报文配置对应的流量监管行为。

（3）配置流策略。创建流策略，关联上面定义的流分类和流行为。

（4）应用流策略。在全局、VLAN或具体的交换机端口上应用上面配置的流策略，进行流量监管。在端口上进行流量监管则此端口单独享有端口限制速率，而在全局进行流量监管则所有端口共享端口限制速率。

说明

相同的流策略可以在不同的接口下应用，当匹配流分类规则的报文的接收或发送 速率超过限制速率时，直接被丢弃。基于流的流量监管，可以通过流分类，为不同业务提供更细致的差分服务。

在S5700SI、S5700LI和S5700S-LI系列交换机中，如果同时配置了接口入方向的流量监管、VLAN的广播流量抑制，以及入方向的基于流的流量监管时，且有报文同时符合上述两种或两种以上限速的条件，限速生效的优先级由高到低依次是接口入方向的流量监管、VLAN的广播流量抑制、入方向的基于流的流量监管。例如，同时匹配了接口入方向的流量监管和VLAN的广播流量抑制，则接口入方向的流量监管生效。

11.2.3 配置流量整形

流量整形（TS）是一种主动调整流量输出速率的措施。当下游设备的入接口速率小于上游设备的出接口速率或发生突发流量时，下游设备入接口处可能出现流量拥塞的情况，此时用户可以通过在上游设备 的出接口上配置流量整形，将上游不规整的流量得到整形，输出一条速率比较平整的流量，从而解决下游设备的拥塞问题。配置流量整形后，可实现报文的流量以均匀的速率向外发送，减少因超过承诺速率而被丢弃的报文。

与流量监管相同，流量整形也是对流量进行限速。但是利用流量监管进行限速时，系统会直接丢弃不符合速率要求的报文，而流量整形将不符合速率要求的报文先送入队列进行缓存，当令牌桶有足够的令牌时，再均匀地向外发送这些被缓存的报文。流量整形会增加延迟（由于流量整形采用了缓存机制），而流量监管几乎不引入额外的延迟。

所有华为S系列交换机均支持以下两种流量整形功能（可同时配置 ）。

● 基于接口的流量整形

对接口上所有通过的报文进行流量整形。

● 基于队列的流量整形

对接口上通过的某类报文（基于简单流分类）分别进行流量整形，从而可实现针对业务（如话音、数据、视频）的流量整形。

1．基于接口的流量整形配置

如果需要对接口出方向所有流量进行控制，可以配置基于接口的流量整形。这样，当接口发送报文的速率超过限制速率时，超出的那部分报文先进入缓存队列；当令牌桶有足够的令牌时，再均匀地向外发送这些被缓存的报文；当缓存队列已满时，报文将被丢弃。

华为S系列交换机配置基于接口流量整形的方法是在对应接口下使用qos lr outbound cir cir-value [cbs cbs-value]（S2700/3700/5700/6700 系列中）命令，或 qos lr cir cir-value [cbs cbs-value] [outbound]（S7700/9300/9300E/9700系列中）命令进行的。如 果多个接口需要配置相同的流量整形速率，可通过接口组进行配置，以减少重复配置工作。命令中的参数说明如下。

（1）cir-value：指定接口承诺信息速率，即保证能够通过的平均速率，单位为kbit/s，但不同的接口类型取值范围不同，如标准以太网接口的取值范围为 64～100 000的整数， GE接口的取值范围为 64～1 000 000的整数，10GE接口的取值范围为 64～10 000 000的整数，40GE接口的取值范围为 64～40 000 000。

（2）cbs-value：可选参数，指定承诺突发尺寸，即瞬间能够通过的承诺突发流量。单位为 byte（字节），取值范围为 4 000～4 294 967 295的整数。若不指定该参数：单令牌桶时，本参数的缺省值为参数cir-value取值的188倍；双令牌桶时，本参数的缺省值为cir-value参数取值的125倍。

本命令为覆盖式命令，即在同一接口多次配置流量整形参数后，按最后一次配置生效 。缺省情况下，接口上不进行流量整形，即整形速率缺省为接口的最大带宽（如Ethernet接口为100 000kbit/s，GE接口为1 000 000kbit/s），可用undo qos lr outbound（S2700/3700系列交换机中）或undo qos lr（S5700/6700/7700/9300/9700系列中）命令用来取消对应接口或者接口组上的流量整形功能。

【示例 1】在S2700/3700/5700/6700系列中，限制GE0/0/1接口向外发送数据的承诺信息速率为 20 000kbit/s，承诺突发尺寸为 375 000字节。

```
<HUAWEI > system-view
[HUAWEI] interface gigabitethernet 0/0/1
[HUAWEI-GigabitEthernet0/0/1] qos lr outbound cir 20000 cbs 375000
[HUAWEI-GigabitEthernet0/0/1] quit
```

【示例 2】在S7700/9300/9300E/9700系列中，限制GE1/0/1接口向外发送数据的承诺信息速率为 20 000kbit/s，承诺突发尺寸为 375 000字节。

```
<HUAWEI> system-view
[HUAWEI] interface gigabitethernet1/0/1
[HUAWEI-GigabitEthernet1/0/1] qos lr cir20000 cbs 375000 outbound
[HUAWEI-GigabitEthernet1/0/1] quit
```

【示例 3】在S7700/9300/9300E/9700系列中，通过接口组group1，限制GE1/0/1～GE1/0/6接口的CIR为200 000kbit/s，CBS为 3 750 000字节。

```
<Qu idway> system-view
[HUAWEI] port-group group1
[HUAWEI-port-group-group1] group-member gigabitethernet 1/0/1 to gigabitethernet 1/0/6
[HUAWEI-port-group-group1] qos lr cir 200000 cbs 3750000
[HUAWEI-port-group-group1] quit
```

2．配置基于接口队列的流量整形

上面介绍的基于接口的流量整形是不区分报文类型的，对所有从该接口发送的报文都进行同样的整形。此处介绍的基于接口队列的流量整形可以实现对不同类型报文区分整形。对从交换机内部进入到出接

口的报文先根据所配置的优先级映射进入到出接口的不同队列，然后针对不同的优先级队列设置不同的流量整形参数，以实现对不同业务的差分服务。也正因如此，在配置基于接口队列的流量整形前，需要配置优先级映射，将报文的优先级映射为PHB行为，从而使不同业务进入不同的接口队列。有关优先级映射的配置参见第10章10.2节。

不同的 S 交换机系列，配置基于接口队列的流量整形的方法有所不同。（1）在除S2700-52P-EI/2700-52P-PWR-EI之外的其他S2700EI系列交换机中，需要在对应的出接口视图下使用 qos queue queue-index shaping cir cir-value cbs cbs-value 命令进行配置。（2）在其他 S 系列交换机中，需要在对应的出接口视图下使用 qos queue queue-index shaping cir cir-valuepirpir-value [cbscbs-valuepbs pbs-value] 命令进行配置。这两个命令中的参数说明如下。

（1）queue-index：指定队列索引号，除S2700系列的取值范围为0～3外，其他所有S系列交换机的取值范围为0～7。可为接口的各个队列配置整形功能。

（2）cir-value：指定接口承诺信息速率，即保证能够通过的平均速率，单位为kbit/s，但不同的接口类型取值范围不同：标准以太网接口的取值范围为 0～100 000的整数，千兆以太网接口的取值范围为0～1 000 000的整数，10GE接口的取值范围为0～10 000 000的整数，40GE接口的取值范围为 0～40 000 000。缺省值为端口的最大带宽。

（3）pir-value：指定峰值信息速率，即最大能够通过的速率。单位为kbit/s，但不同的接口类型取值范围不同：标准以太网接口的取值范围为 64～100 000的整数，GE接口的取值范围为 64～1 000 000的整数，10GE接口的取值范围为 64～10 000 000的整数， 40GE 接口的取值范围为 64～40 000 000。缺省值为端口的最大带宽，且必须大于等于cir-value参数值。

（4）cbs-value：可选参数，指定承诺突发尺寸，即瞬间能够通过的承诺突发流量。单位为 byte（字节），取值范围为 4 000～4 294 967 295（S2700/3700/5700/6700系列中）或 10 000～4 294 967 295（S7700/9300/9300E/9700系列中）的整数。若不指定该参数：单令牌桶时，本参数的缺省值为参数cir-value取值的188倍；双令牌桶时，本参数的缺省值为cir-value参数取值的125倍。

（5）pbs-value：指定峰值突发尺寸，即瞬间能够通过的峰值突发流量。整数形式，取值范围是 10 000～4 294 967 295，单位是 byte。缺省值与配置的 pir-value有关

均可用undo qos queue queue-index shaping命令恢复接口上各参数的缺省值。

【示例 4】在S2700-52P-EI/2700-52P-PWR-EI系列交换机GE0/0/1端口2队列上配置队列整形功能，参数CIR为 1 000kbit/s、CBS为 125 000byte。

```
<HUAWEI> system view
[HUAWEI] interface gigabitethernet 0/0/1
[HUAWEI-GigabitEthernet0/0/1] qos queue 2 shaping cir 1000 cbs 125000
[HUAWEI-GigabitEthernet0/0/1] quit
```

【示例 5】在S2700-52P-EI/2700-52P-PWR-EI系列之外的S系列交换机GE1/0/1端口4队列上配置队列整形功能，参数CIR为 10 000kbit/s、PIR为 20 000kbit/s。

```
<HUAWEI> system-view
[HUAWEI] interfacegigabitethernet 1/0/1
[HUAWEI-GigabitEthernet1/0/1] qosqueue 4 shaping cir 10000 pir 20000
[HUAWEI-GigabitEthernet1/0/1] quit
```

说明

如果在同一接口下既配置接口队列整形功能，又配置接口整形功能，则接口整形 的CIR必须大于等于

接口所有队列整形的CIR之和；否则，流量整形会出现异常现象，如低优先级队列抢占高优先级队列的带宽等。

3．（可选）配置接口队列缓存

可以任一流量整形功能配置接口队列的缓存。缓存大小会影响到队列整形的效果，可以根据网络实际需求，配置接口指定队列缓存大小。修改接口队列缓存前需要使用shutdown命令关闭接口，配置完后再使用undo shutdown打开接口，否则配置不生效。同样不同S系列交换机的配置方法有所不同。

（1）在S2700/3700/5700EI系列交换机中

在具体接口或者接口组下使用qos queue queue-indexmax-length packet-number命令配置接口队列缓存大小，但只有S2700-52P-EI、S2700-52P-PWR-EI、S2710SI、S3700SI、S3700EI支持此命令。命令中的参数说明如下。

① queue-index：指定队列索引号，除S2700系列的取值范围为0～3外，其他所有S系列交换机的取值范围为0～7的整数。可为接口的各个队列配置缓存大小。

②packet-number：指定该队列最大可缓存的报文个数，取值范围是5～980的整数。

缺省情况下，队列0可以缓存报文的最大报文个数是400，其他队列可以缓存报文的最大报文个数是89，可用undo qos queue queue-indexmax-length命令用来恢复对应队列可以缓存报文的最大报文个数到缺省值。

【示例 6】在S2700/3700/5700EI系列交换机中配置GigabitEthernet0/0/1接口队列0可以缓存报文的最大报文个数为155。

```
<HUAWEI> system-view
[HUAWEI] interface gigabitethernet 0/0/1
[HUAWEI-GigabitEthernet0/0/1] qos queue 0max-length 155
```

（2）在S5700HI/6700/7700/9300/9300E/9700系列交换机中

在具体接口或者接口组下使用qos queue queue-index length length-value命令配置接口的优先级队列长度。命令中的参数说明如下。

① queue-index：指定队列索引号，取值范围为0～7的整数。可为接口的各个队列配置缓存大小。

② length-value：指定该队列的最大长度，单位为 byte，取值范围是 0～1 000 000 000的整数。

缺省情况下，系统统一自动管理各接口的优先级队列长度，可用 undo qos queue queue-index length命令取消在对应队列上配置的优先级队列长度。

【示例 7】在S5700HI/6700/7700/9300/9300E/9700系列交换机中，配置GE1/0/0接口上的优先级队列 1的长度为 20 000字节。

```
<HUAWEI> system view
[HUAWEI] interface gigabitethernet 1/0/0
[HUAWEI-GigabitEthernet1/0/0] qos queue1 length 20000
```

（3）在S5700SI/5700LI/5700S-LI系列交换机中

在S5700SI/5700LI/5700S-LI系列交换机中，在拥塞发生期间，可以对报文采用尾部丢弃（Tail-Drop）的方法。当队列的长度达到最大值后，所有新入队列的报文（缓存在队列尾部）都将被丢弃，直到拥塞解决。可使用qos queue max-length命令配置接口指定队列可以缓存报文的最大报文个数，确保该队列有足够可用的缓冲区，避免队列因为不能得到缓冲区而丢失流量。采用尾部丢弃法的接口队列缓存的配置比较复杂，具体步 骤如下。

（1）先在系统视图下使用 qos tail-drop-profileprofile-name命令创建全局尾丢弃模板，并进入尾丢弃模

板视图。参数用来profile-name用来指定尾丢弃模板名称，为1～16个字符，不支持空格，不区分大小写，但定义的模板数量不能超过 7 个，否则系统会报错。

缺省情况下，系统没有创建任何全局尾丢弃模板，可用 undo qos tail-drop-profile profile-name命令删除已存在的指定尾丢弃模板。

（2）然后在 5700SI/5700-28P-LI/5700-52P-LI系列交换机中使用 qos queue queue- indexmax-lengthpacket-number [green max-length packet-number] 命令，在S5700-10P-LI/ 5700-28X-LI/5700-52X-LI 系列交换机中使用 qos queue queue-indexgreen max-length packet-numbernon-green max-length packet-number命令配置接口队列缓存大小。两命令中的参数说明如下。

① queue-index：指定队列索引号，取值范围为 0～7。可为接口的各个队列配置缓存大小。

②max-length packet-number：指定队列中可以缓存的最大报文个数，S5700SI系列交换机的取值范围为1～5 134的整数，各 S5700LI和 S5700-LI系列交换机的取值范围是 1～3 000的整数。

③ green max-length packet-number：指定队列中可以缓存绿色报文的最大个数， S5700SI 系列交换机的取值范围为 1～5 134 的整数，S5700-28P-LI、S5700-52P-LI 和S5700S-LI系列交换机的取值范围为 1～3 000的整数，S5700-10P-LI、S5700-28X-LI和S5700-52X-LI系列交换机的取值范围为 1 280～3 000的整数。

④ non-green max-lengthpacket-number：指定队列中可以缓存非绿色报文的最大个数，取值范围为 1 280～3 000的整数。

缺省情况下，S5700SI、S5700-28P-LI、S5700-52P-LI队列的全部报文最大缓存为22，绿色报文最大缓存为 11，可用undo qos queue queue-indexmax-length [packet-number green max-length packet-number | green max-length] 命令恢复队列可以缓存绿色报文的最大个数为缺省值。

缺省情况下，S5700-10P-LI、S5700-28X-LI、S5700-52X-LI 队列绿色报文最大缓存为 1280，非绿色报文最大缓存为 1 280，可用undo qos queue queue-index green max- length [packet-number] non-green max-length [packet-number] 命令恢复队列可以缓存绿色报文的最大个数为缺省值。

【示例 8】全局名称为test的尾丢弃模板，配置队列0可以缓存的最大报文个数为200。

```
<HUAWEI>system-view
[HUAWEI] qos tail-drop-profile test
[HUAWEI-tail-drop-profile-test] qos queue 0 max-length 200
```

（3）最后在关闭接口或者接口组的情况下，进入对应的接口或接口组视图使用 qos tail-drop-profile profile-name命令在接口或者接口组下应用指定的尾丢弃模板。

【示例 9】全局创建名称为test的尾丢弃模板，配置队列1的绿色报文最大长度为10，并在GE0/0/1接口下应用该模板。

```
<HUAWEI>system-view
[HUAWEI] qos tail-drop-profile test
[HUAWEI-tail-drop-profile-test] qos queue 1 green max-length 10
[HUAWEI-tail-drop-profile-test] quit
[HUAWEI] interfacegigabitethernet 0/0/1
[HUAWEI-GigabitEthernet0/0/1] qos tail-drop-profile test
```

11.2.4 流量监管和流量整形管理

在完成流量监管和流量整形配置后，可以根据不同机型选择对应的 display 任意视图命令查看流量监管和流量整形配置，或接口或队列上的流量统计信息。也可使用以下reset用户视图命令清除流量统计数据。

因篇幅限制，在此就不具体举例了。

1．在S2700/3700系列交换机中

（1）执行display qos lr outbound interface interface-type interface-number命令可查看指定接口上的速率限制信息。

（2）执行 reset traffic policy statistics {global [slot slot-id] | interface interface-type interface-number | vlan vlan-id }inbound命令可清除全局、指定接口或指定VLAN下应用的流策略的统计信息。

（3）在 S2700-52P-EI/2700-52P-PWR-EI/3700SI/3700EI系列交换机上执行 reset qos queue statistics [queue queue-index { inbound interfaceinterface-type interface-number |outbound interface interface-type interface-number [from interface { interface-type interface-number| all }] }] 命令可清除接口上基于队列的流量统计信息。

2．在S5700/6700系列交换机中

【说明】在S5700SI和S5700EI系列交换机中查看接口队列的整形参数之前，需要先执行qos queue statistics enable命令使能对指定接口队列的统计功能。

（1）执行 display traffic policy statistics {global [slot slot-id] | interface interface- type interface-number | vlan vlan-id } {inbound | outbound } [verbose { classifier-base |rule-base } [class classifier-name]] 命令可查看基于流的流量统计信息。

（2）在S5700SI系列交换机中，执行display qos queue statistics [queue queue-index outbound interface interface-type interface-number] 命令可查看接口队列的统计信息。

（3）在S5700EI系列交换机中，执行display qos queue statistics [queue queue-index { inbound interface interface-type interface-number | outbound interface interface-type interface-number [from interface { interface-type interface-number | all }] }] 命令可查看接口队列的统计信息。

（4）在S5700HI/5710EI/5700LI/5700S-LI/6700系列交换机上执行display qos queue statistics interface interface-type interface-number命令可查看接口队列的统计信息。

（5）在S5700SI系列交换机中，执行reset qos queue statistics [queuequeue-indexoutbound interface interface-type interface-number] 命令可清除接口上基于队列的流量统计信息。

（6）在S5700EI系列交换机中，执行 reset qos queue statistics [queue queue-index { inbound interface interface-type interface-number | outbound interface interface-type interface-number [from interface { interface-type interface-number | all }] }] 命令可清除接口上基于队列的流量统计信息。

（7）在 S5700HI/5700LI/5700S-LI/5710EI/6700 系列交换机上执行 reset qos queue statistics interface interface-type interface-number命令清除接口上基于队列的流量统计信息。

3．在S7700/9300/9300E/9700系列交换机中

（1）执行display traffic policy statistics {global [slot slot-id] | interface interface-type interface-number |vlanvlan-id } {inbound |outbound } [verbose {classifier-base |rule-base } [class classifier-name]] 命令可查看基于流的流量统计信息。

（2）执行display qos car statistics interface interface-type interface-number inbound命令可查看配置基于接口的流量监管后指定接口上通过和丢弃的报文统计信息。

（3）执行display qos queue statistics interface interface-type interface-number命令可查看接口上基于队列的流量统计信息。

（4）执行 reset qos car statistics interface interface-type interface-number inbound命令可清除配置基于接口的流量监管后指定接口上通过和丢弃的报文统计信息。

（5）执行 reset qos queue statistics interface interface-type interface-number命令可清除接口上基于队列的流量统计信息。

11.2.5 基于接口的流量监管配置示例

本示例拓扑结构如图11-3所示，Switch通过GE0/0/3接口与路由器互连，企业分支机构1和企业分支机构2通过GE0/0/1和GE0/0/2接口接入Switch，经由Switch和路由器访问网络。现要求企业分支机构1入方向保证带宽为8Mbit/s，企业分支机构2入方向保证带宽为5Mbit/s。

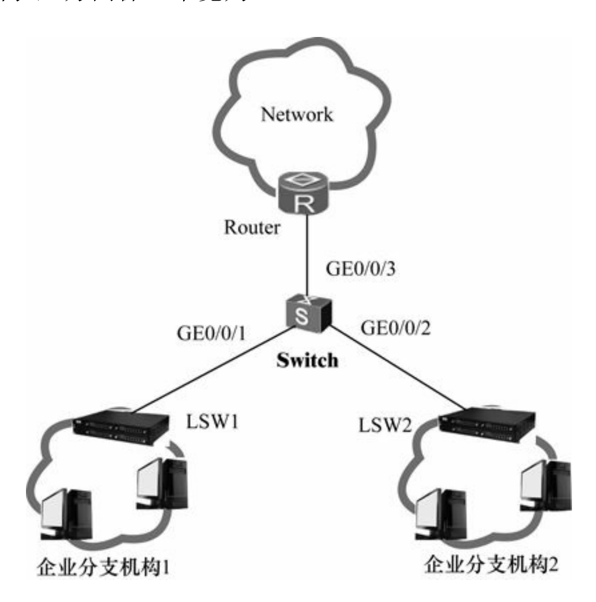


图11-3 基于接口的流量监管配置示例

本示例的配置方法很简单，就是在Switch的GE0/0/1和GE0/0/2接口入方向配置流量监管。因为在流量监管配置方法上S2700/3700/5700/6700系列交换机与S7700/9300/9300E/9700系列是完全不同的，所以下面分别介绍当Switch设备为以上两种情形下的具体配置步骤。

1．在S2700/3700/5700/6700系列交换机上的具体配置步骤

S2700/3700/5700/6700 系列交换机上的流量监管是在具体接口上配置的，具体配置 步骤如下。

（1）在GE0/0/1接口入方向上配置流量监管，保证带宽为 8 192kbit/s（由8Mbit/s转换得到）。

```
[Switch] interface gigabitethernet 0/0/1
[Switch-GigabitEthernet0/0/1] qos lr inbound cir 8192
[Switch-GigabitEthernet0/0/1] quit
```

（2）在GE0/0/2接口入方向上配置流量监管，保证带宽为 5 120kbit/s。

```
[Switch] interface gigabitethernet 0/0/2
[Switch-GigabitEthernet0/0/2] qos lr inbound cir5120
[Switch-GigabitEthernet0/0/2] quit
```

配置好后，分别利用display qos lr inbound interface gigabitethernet 0/0/1和display qos lr inbound interface gigabitethernet 0/0/2命令查看两接口上的流量监管配置信息，验证流量监管配置结果，具体如下，从输出信息中的粗体字部分可以看出配置是正确的（cbs值是按照双令牌桶时缺省情况为125倍cir值得到的）。

```
[Switch] display qos lr inbound interfacegigabitethernet 0/0/1
```

GigabitEthernet0/0/1 lr inbound:

cir: 8192 Kbps, cbs: 1024000 Byte

[Switch] display qos lr inbound interfacegigabitethernet0/0/2

GigabitEthernet0/0/2 lr inbound:

cir: 5120 Kbps, cbs: 640000 Byte

2．在S7700/9300/9300E/9700系列交换机上的具体配置步骤

S7700/9300/9300E/9700系列交换机中的流量监管是先要创建对应的CAR模板，然后在具体接口视图（或者具体VLAN视图、流行为视图）下应用这个CAR模板的。具体配置步骤如下。

（1）在Switch上创建CAR模板car1、car2，分别对企业分支机构1和企业分支机构2的流量进行监管。

[Switch] qos car car1 cir 8192pir 10240

[Switch] qos car car2 cir 5120 pir 8192

（2）在Switch的GE1/0/1、GE1/0/2接口入方向上分别应用car1、car2，对企业分支机构1和企业分支机构2的流量进行监管。

[Switch] interface gigabitethernet1/0/1

[Switch-GigabitEthernet1/0/1] qos car inbound car1

[Switch-GigabitEthernet1/0/1] quit

[Switch] interface gigabitethernet1/0/2

[Switch-GigabitEthernet1/0/2] qos car inbound car2

[Switch-GigabitEthernet1/0/2] quit

配置好后，可以通过display qos car all命令查看CAR模板的配置信息，验证配置结果。具体如下，从输出信息中的粗体字部分也可以看出两接口上应用的流量监管配置是正确的。

<Switch>display qos car all

CAR Name : car1

CAR Index : 0

car cir 8192 (Kbps) pir 10240 (Kbps) cbs 1024000 (byte) pbs 1280000 (byte)

CAR Name : car2

CAR Index : 1

car cir 5120 (Kbps) pir 8192 (Kbps) cbs 640000 (byte) pbs 1024000 (byte)

11.2.6 流量整形配置示例

本示例拓扑结构如图11-4所示，Switch通过GE0/0/2接口与路由器互连，来自Internet的业务有话音、视频、数据，携带的 802.1p优先级分别为 6、5、2（假设均在VLAN 10中），这些业务可经由路由器和Switch到达用户。由于来自网络侧的流量速率大于LSW设备入接口的速率，所以在Switch的出接口GE0/0/1处可能会发生带宽抖动。

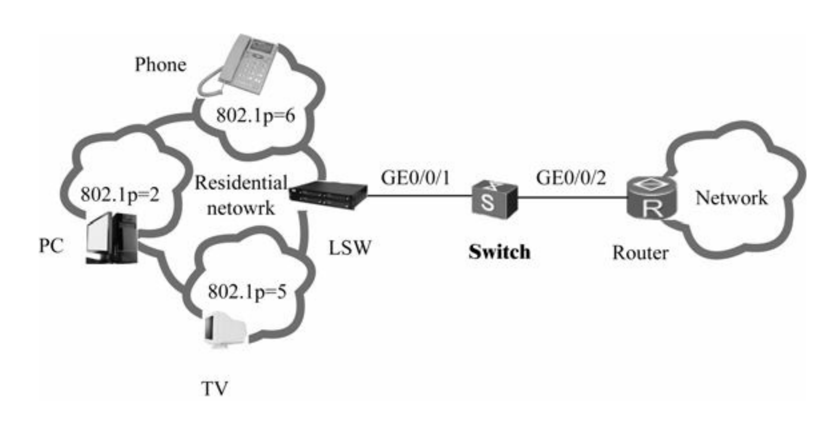


图11-4 流量整形配置示例拓扑结构

现要求在GE0/0/1出接口上做如下配置。

● 保证带宽为 10 000kbit/s。

● 为话音流量保证带宽 3 000kbit/s，峰值带宽 5 000kbit/s。

● 为视频流量保证带宽 5 000kbit/s，峰值带宽 8 000kbit/s。

● 为数据流量保证带宽为 2 000kbit/s，峰值带宽 3 000kbit/s。

1．基本配置思路分析

以下配置均是在Switch上进行的。

（1）创建VLAN 10，并配置各接口加入VLAN 10中，使用户能够通过Switch访问网络。

（2）配置GE0/0/1接口信任报文的802.1p优先级。

（3）配置基于GE0/0/1接口的流量整形功能，限制端口带宽为 10 000kbit/s。

（4）配置端口队列整形功能，限制话音、视频、数据三类业务符合示例中要求的带宽。

2．具体配置步骤

（1）创建VLAN 10。

```
<HUAWEI>system-view
[HUAWEI] sysname Switch
[Switch] vlan batch 10
```

（2）配置GE0/0/1、GE0/0/2接口类型均为 trunk类型，并将它们都加入VLAN 10中。

```
[Switch] interface gigabitethernet 0/0/1
[Switch-GigabitEthernet0/0/1] port link-type trunk
[Switch-GigabitEthernet0/0/1] port trunk allow-pass vlan 10
[Switch-GigabitEthernet0/0/1] quit
[Switch] interfacegigabitethernet 0/0/2
[Switch-GigabitEthernet0/0/2] port link-type trunk
[Switch-GigabitEthernet0/0/2] port trunk allow-pass vlan 10
[Switch-GigabitEthernet0/0/2] quit
```

（3）创建VLANIF10，并配置网段地址10.10.10.1/24。这样做的目的是通过VLANIF10接口与 Router 之间建立三层连接，当然需要在 Router 与 Switch 连接的接口上配置与VLANIF10同网段的IP地址，如10.10.10.2/24。

```
[Switch] interface vlanif 10
```

[Switch-Vlanif10] ip address 10.10.10.1 255.255.255.0

[Switch-Vlanif10] quit

（4）配置GE0/0/1接口信任报文的802.1p优先级。

[Switch] interface gigabitethernet 0/0/1

[Switch-GigabitEthernet0/0/1] trust 8021p

[Switch-GigabitEthernet0/0/1] quit

（5）配置基于GE0/0/1接口的流量整形，将端口速率限制在 10 000kbit/s。

[Switch] interfacegigabitethernet 0/0/1

[Switch-GigabitEthernet0/0/1] qos lr outbound cir10000

（6）在Switch的GE0/0/1接口上配置端口队列整形，使话音、视频、数据业务的保证带宽分别为 3 000kbit/s、5 000kbit/s、2 000kbit/s，峰值带宽分别为 5 000kbit/s、8 000kbit/s、3 000kbit/s。

说明

这里没有配置优先级与队列之间的映射，而是直接采用缺省的802.1p优先级与队 列之间的一一对应映射，即 0～7 802.1p优先级分别对应 0～7号队列。

[Switch-GigabitEthernet0/0/1] qos queue 6 shaping cir 3000 pir5000

[Switch-GigabitEthernet0/0/1] qos queue 5 shaping cir 5000 pir8000

[Switch-GigabitEthernet0/0/1] qos queue 2 shaping cir 2000 pir3000

[Switch-GigabitEthernet0/0/1] quit

配置成功后，从GE0/0/1接口发出的报文保证速率不超过 10 000kbit/s；话音业务保证速率为 3 000kbit/s，不超过 5 000kbit/s；视频业务保证速率为 5 000kbit/s，不超过8 000kbit/s；数据业务保证速率为 2 000kbit/s，不超过 3 000kbit/s。

11.3 拥塞避免和拥塞管理的配置与管理

拥塞避免通过指定报文丢弃策略来解除网络过载，拥塞管理通过指定报文的调度次序来确保高优先级业务优先被处理。各S系列交换机所支持的拥塞避免和拥塞管理功能参见第10章10.4节。

注意

在对拥塞避免，以及RED、SRED和WRED技术的支持上注意以下几个方面。

（1）S2700SI和S2700EI系列不支持拥塞避免功能。

（2）S5700SI/5700LI/5700S-LI系列仅支持尾部丢弃拥塞避免方法。

（3）S2700-52P-EI/2700-52P-PWR-EI/2710SI/3700SI/3700EI/5700EI系列仅支持SRED技术。

（4）S5700HI/5710EI/6700/7700/9300/9300E/9700系列仅支持WRED技术。

11.3.1 尾部丢弃法拥塞避免的配置与管理

S5700SI/5700LI/5700S-LI 系列交换机仅支持尾部丢弃的方法实现拥塞避免，当队列的 长度达到最大值后，所有新入队列的报文（缓存在队列尾部）都将被丢弃。通过增加端口队列的缓存大小，可以避免报文因为不能得到缓存而丢失流量。具体配置步骤如表11-11所示。

表11-11 尾部丢弃法拥塞避免配置步骤

步骤	命令	说明
1	system-view 例如：<HUAWEI> system-view	进入系统视图
2	**qos tail-drop-profile** *profile-name* 例如：[HUAWEI] **qos tail-drop-profile** test	创建尾丢弃模板，并进入尾丢弃模板视图。参数用来 *profile-name* 用来指定尾丢弃模板名称，为 1～16 个字符，不支持空格，不区分大小写，但定义的模板数量不能超过 7 个，否则系统会报错
3	在 S5700SI/5700-28P-LI/5700-52P-LI 系列交换机中： **qos queue** *queue-index* **max-buffer** *cell-number* [**green max-buffer** *cell-number*] 或 **qos queue** *queue-index* **green max-buffer** *cell-number* 例如：[HUAWEI-tail-drop-profile-test] **qos queue** 0 **max-buffer** 100 在 S5700-10P-LI/5700-28X-LI/5700-52X-LI 系列交换机中： **qos queue** *queue-index* **green max-buffer** *cell-number* **non-green max-buffer***cell-number* 例如：[HUAWEI-tail-drop-profile-test] **qos queue** 0 **green max-buffer** 100 **non-green max-buffer** 10	配置指定队列的全部报文的最大缓存或绿色报文的最大缓存。命令中的参数说明如下。 （1）*queue-index*：指定队列索引号，取值范围为 0~7 的整数 （2）**max-buffer** *cell-number*：指定队列中全部报文的最大缓存字节数，单位是 cell，一个 cell 的大小为 128 个字节。不同系列的取值范围不一样：S5700SI 系列取值范围为 1～5 444 的整数，S5700-28P-LI/5700-52P-LI/5700S-LI 系列的取值范围为 1～3 100 的整数，S5700-10P-LI/5700-28X-LI/5700-52X-LI 系列的取值范围为 1 920～3 100 的整数 （3）**green max-buffer** *cell-number*：指定队列中绿色报文的最大缓存字节数，单位也是 cell，取值范围同 **max-buffer** *cell-number* （4）**non-green max-buffer** *cell-number*：指定队列中可以缓存非绿色报文的最大字节数，单位也是 cell，取值范围为 1 920～3 100 的整数 缺省情况下，S5700SI/5700-28P-LI/5700-52P-LI 系列交换机队列的全部报文最大缓存为 24，绿色报文最大缓存为 12，单位是 cell；缺省情况下，S5700-10P-LI/5700-28X-LI/5700-52X-LI 系列交换机队列绿色报文最大缓存为 1 280，非绿色报文最大缓存为 1 280，单位是 cell，可用对应的 **undo** 格式命令进行恢复
4	在 S5700SI/5700-28P-LI/5700-52P-LI 系列交换机中： **qos queue** *queue-index* **max-length** *packet-number* [**green max-length** *packet-number*] 或 **qos queue** *queue-index* **green max-length** *packet-number* 例如：[HUAWEI-tail-drop-profile-test]**qos queue** 0 **max-length** 200 在 S5700-10P-LI/5700-28X-LI/5700- 52X-LI 系列交换机中： **qos queue green max-length** *packet-number* **non-green max-length** *packet-number* 例如：[HUAWEI-tail-drop-profile-test]**qos queue** 0 **max-length** 200 **non-green max-length** 10	配置指定队列可以缓存的全部报文最大数或者可以缓存的绿色报文的最大数。命令中的参数说明如下。 （1）*queue-index*：指定队列索引号，取值范围为 0~7 的整数 （2）**max-length** *packet-number*：指定队列中可以缓存全部报文的最大个数。S5700SI 系列的取值范围为 1~5 134 的整数，S5700LI/5700S-LI 系列的取值范围为 1~3 000 （3）**green max-length** *packet-number*：指定队列中可以缓存绿色报文的最大个数。S5700SI 系列的取值范围是 1～5 134 的整数，S5700-28P-LI/5700-52P-LI/5700S-LI 系列的取值范围为 1～3 000 的整数，S5700-10P-LI/5700-28X-LI/5700-52X-LI 系列的取值范围为 1 280～3 000 的整数 （4）**non-green max-length** *packet-number*：指定队列中可以缓存非绿色报文的最大个数，取值范围为 1 280～3 000 的整数 缺省情况下，S5700SI/5700-28P-LI/5700-52P-LI 系列交换机队列的全部报文最大缓存为 22，绿色报文最大缓存为 11；S5700-10P-LI/5700-28X-LI/5700-52X-LI 系列交换机队列绿色报文最大缓存为 1 280，非绿色报文最大缓存为 1 280，可用对应的 **undo** 格式命令进行恢复

（续表）

步骤	命令	说明
5	**quit** 例如：[HUAWEI-tail-drop-profile-test] **quit**	退出尾丢弃模板视图，返回系统视图
6	**Interface** *interface-type interface-number* 例如：[HUAWEI] **interface** gigabitethernet 0/0/1	键入要配置尾部丢弃方法拥塞避免功能的接口，进入接口视图
7	**shutdown** 例如：[HUAWEI-GigabitEthernet0/0/1] **shutdown**	关闭以上接口
8	**qos tail-drop-profile** *profile-name* 例如：[HUAWEI-GigabitEthernet0/0/1] **qos tail-drop-profile** test	在接口下应用以上配置的尾丢弃模板（**必须在接口关闭状态下应用**）。缺省情况下，接口下没有应用任何尾丢弃模板，可用 **undo qos tail-drop-profile** 命令删除在接口下应用的尾丢弃模板
9	**undo shutdown** 例如：[HUAWEI-GigabitEthernet0/0/1] **undo shutdown**	打开接口（**应用尾丢弃模板后必须重新开启接口，使配置生效**）

【示例 1】在全局下创建模板名称为test的尾丢弃模板，配置队列0的最大报文缓存为100。

<HUAWEI> system-view

[HUAWEI] qos tail-drop-profile test

[HUAWEI-tail-drop-profile-test] qos queue 0max-buffer100

【示例 2】在全局下创建模板名称为test的尾丢弃模板，配置队列1的绿色报文最大长度为10，并在GE0/0/1接口下应用该模板。

```
<HUAWEI>system-view
[HUAWEI] qos tail-drop-profile test
[HUAWEI-tail-drop-profile-test] qos queue 1 green max-length 10
[HUAWEI-tail-drop-profile-test] quit
[HUAWEI] interface gigabitethernet 0/0/1
[HUAWEI-GigabitEthernet0/0/1] shutdown
[HUAWEI-GigabitEthernet0/0/1] qos tail-drop-profile test
[HUAWEI-GigabitEthernet0/0/1] undo shutdown
```

配置好后可在任意视图下执行 display qos configuration interface interface-type interface-number命令查看接口上所有的QoS配置信息。在S5700LI/5700SI/5700S-LI系列交换机上执行display qos queue statistics interface interface-type interface-number任意视图命令查看接口上基于队列的流量统计信息；执行display qos queue statistics [queue queue-index outbound interface interface-type interface-number] 任意视图命令查看接口上基于队列的流量统计信息。

11.3.2 SRED拥塞避免的配置与管理

在S2700-52P-EI/2700-52P-PWR-EI/2710SI/3700SI/3700EI/5700EI系列交换机仅支持SRED方法的拥塞避免功能。设备根据SRED的配置信息对不同颜色的报文按照一定的丢弃概率主动丢弃队列中的报文，从而调整从接口输出的流量速率。

在配置 SRED 拥塞避免功能之前，需在报文的入接口上完成以下任务之一为报文着色 （当然，也可直接采用缺省的优先级与**PHB**行为/颜色映射配置 ），作为拥塞避免操作的依据。

● 配置基于ACL的简单流分类的流量监管，将报文的优先级映射为PHB行为并着色，参见本书第9章9.6.3节。

● 配置基于复杂流分类的流量监管和重标记，具体将在本章后面11.4节介绍。

SRED拥塞避免功能的配置任务包括以下3个方面。

1．（可选）配置接口队列缓存

配置接口队列的缓存大小，确保该队列有足够可用的缓冲区，可以避免报文因为不能得到缓存而丢失流量。但在接口上配置接口队列缓存前需要使用shutdown命令关闭接口，配置完成后，再使用undo shutdown命令打开接口，此操作过程可能会引起网络的短暂中断。

2．（可选）配置CFI作为内部丢弃优先级

VLAN标签中的CFI（Canonical Format Indicator，规范格式指示器）字段又称为DEI （Drop Eligible Indicator，丢弃资格指示器）。在某些情况下用来标识报文的丢弃优先级，某些设备在报文超出CIR（承诺信息速率）时会将报文的DEI位置1，标识该报文的丢弃优先级为高，后续设备在拥塞的时候优先丢弃DEI位为1的报文。如果多个接口需要配置CFI作为内部丢弃优先级，可通过端口组进行配置，以减少重复配置工作。

3．配置SRED参数

在接口出队列上，SRED 根据报文的优先级将其区分为红色、黄色，并分别为红色和黄色的报文设定起始丢包点和丢包率，从而实现拥塞避免。

在配置基于SRED的拥塞避免时，对0～4队列设置红色的丢包点和丢弃率起作用，设置黄色的丢包点和

丢包率不起作用；对5～7队列设置黄色的丢包点和丢包率起作用，设置红色的丢包点和丢包率不起作用。

以上配置任务的具体配置步骤如表11-12所示。

表11-12 SRED方法拥塞避免配置步骤

配置任务	步骤	命令	说明
（可选）配置端口队列长度	1	**system-view** 例如：<HUAWEI> **system-view**	进入系统视图
	2	**interface** *interface-type interface-number* 例如：[HUAWEI] **interface** gigabitethernet 0/0/1	键入要配置 SRED 方法拥塞避免功能的接口，进入接口视图
	3	**shutdown** 例如：[HUAWEI-GigabitEthernet0/0/1] **shutdown**	关闭以上接口
	4	**qos queue** *queue-index* **static-cell** *cell-number* 例如：[HUAWEI-GigabitEthernet0/0/1] **qos queue** 1 **static-cell** 1000	配置接口指定队列的静态缓冲区字节数。命令中的参数说明如下。 （1）*queue-index*：指定队列的索引号，取值范围 0～7 的整数 （2）*cell-number*：指定静态缓冲区的大小，取值范围为 1～16 140 的整数，单位是 cell，一个 cell 的大小为 128 字节。缺省值是 27，可用 **undo qos queue** *queue-index* **static-cell** 命令恢复接口指定队列静态缓冲区字节数为缺省值

（续表）

配置任务	步骤	命令	说明
（可选）配置端口队列长度	5	**qos queue** *queue-index* **max-length** *packet-number* 例如：[HUAWEI-GigabitEthernet0/0/1] **qos** queue 1 max-length 100	配置队列最大缓存的报文数。上步配置的最大字节数和本步配置的最大报文数只要有一个被占满，则认为网络发生拥塞，后续报文开始丢弃。命令中的参数说明如下。 （1）*queue-index*：指定队列的索引号，S3700/5700 系列的取值范围 0～7，S2700 系列的取值范围为 0～3 （2）*packet-number*：配置接口指定队列可以缓存报文的最大报文个数，取值范围为 5～2 007 缺省情况下，队列 0 可以缓存报文的最大报文个数是 1 044，其他队列可以缓存报文的最大报文个数是 143，可用 **undo qos queue** *queue-index* **max-length** 命令恢复队列可以缓存报文的最大报文个数为缺省值
	6	**undo shutdown** 例如：[HUAWEI-GigabitEthernet0/0/1] **undo shutdown**	打开接口
（可选）配置 CFI 作为内部丢弃优先级	7	**dei enable** 例如：[HUAWEI-GigabitEthernet0/0/1] **dei enable**	配置 CFI 作为内部丢弃优先级。缺省情况下，VLAN 中 DEI 字段映射为丢弃优先级的功能未使能，可用 **undo dei enable** 命令去使能将 VLAN 中的 DEI 字段映射为丢弃优先级功能
	8	**quit** 例如：[HUAWEI-GigabitEthernet0/0/1] **quit**	退出接口视图，返回系统视图
配置 SRED 参数	9	**qos sred queue** *queue-index* **red** *start-discard-point* **discard-probability** *discard-probability* **yellow** *start-discard-point* **discard-probability** *discard-probability* 例如：[HUAWEI]**qos sred queue** 1 **red** 10 **discard-probability** 5 **yellow** 20 **discard-probability** 4	配置队列的 SRED 阈值和丢弃概率。命令中的参数说明如下。 （1）*queue-index*：指定队列的索引号 （2）*start-discard-point*：表示开始丢弃报文的报文序号（即起始丢包点），取值范围为 4～2 047 的整数 （3）*discard-probability*：表示报文丢弃的概率，取值范围是 0～7 的整数，分别对应以下百分比。 0：100% 1：6.25% 2：3.125% 3：1.562 5% 4：0.781 25% 5：0.390 625% 6：0.195 312 5% 7：0.097 656 25% 每个队列都有各自的拥塞避免参数，因此需要针对每个队列重复执行本步骤 缺省情况下，出方向队列上没有配置 SRED 阈值和丢弃概率，可用 **undo qos sred queue** *queue-index* 命令取消出方向队列上配置的 SRED 阈值和丢弃概率

【示例 1】配置GE0/0/1接口队列0的静态缓冲区大小为15。

<HUAWEI> system-view

[HUAWEI] interfacegigabitethernet 0/0/1

[HUAWEI-GigabitEthernet0/0/1] qos queue 0 static-cell 15

【示例 2】配置接口GE0/0/1队列0可以缓存报文的最大报文个数为155。

<HUAWEI> system-view

[HUAWEI] interface gigabitethernet 0/0/1

[HUAWEI-GigabitEthernet0/0/1] qos queue 0max-length 155

【示例 3】在系统视图下配置队列号为 0 的红色报文的起始丢包点为 10，丢弃概率为5；黄色报文的起始丢包点为20，丢弃概率为4。

<HUAWEI>system-view

[HUAWEI] qos sred queue 0 red 10 discard-probability 5 yellow 20discard-probability 4

配置好后，可在任意视图下执行display qos sred命令查看队列上红色和黄色报文的起始丢包点和丢弃概率；执行display qos configuration interface interface-type interface-number命令（本命令已在上节介绍）查看接口上所有的QoS配置信息。

【示例 4】显示设备当前接口出方向队列的全局SRED配置。

<HUAWEI> display qos sred

Current sred configuration:

qos sred queue-index 1 red 10 discard-probability 2 yellow 10 discard-probability 1

qos sred queue-index 2 red 10 discard-probability 2 yellow 10 discard-probability 1

11.3.3 WRED拥塞避免的配置与管理

在S5700HI/5710EI/6700/7700/9300/9300E/9700系列交换机中支持WRED方法拥塞避免功能。配置好后，设备将根据WRED的配置信息对不同颜色的报文进行相应的处理。但在配置拥塞避免前需在报文的入接口上将报文的优先级映射为**PHB**行为（参见**10.5.6**节），作为拥塞避免操作的依据（当然，也可直接采用缺省的优先级与**PHB**行为/颜色映射配置）。且WRED方法拥塞避免功能只对已知单播流量生效。

WRED方法拥塞避免功能的配置方法与上节介绍的SRED方法拥塞避免功能的配置方法有些相似，也包括以下3项配置任务。

1．（可选）配置端口队列长度

通过配置接口队列的缓存大小，确保该队列有足够可用的缓冲区，可以避免报文因为不能得到缓存而丢失流量。但在接口上配置接口队列缓存前需要使用 shutdown 命令关闭接口，配置完成后，再使用undo shutdown命令打开接口，此操作过程可能会引起网络的短暂中断。

2．（可选）配置CFI作为内部丢弃优先级

VLAN标签中的CFI字段在某些情况下用来标识报文的丢弃优先级，某些设备在报文超出CIR（承诺信息速率）时会将报文的DEI位置1，标识该报文的丢弃优先级为高，后续设备在拥塞的时候优先丢弃DEI位为1的报文。如果多个接口需要配置CFI作为内部丢弃优先级，可通过端口组进行配置，以减少重复配置工作。

3．配置WRED丢弃模板

WRED 技术也是通过随机丢弃报文来避免 TCP 的全局同步现象，它通过报文的不同颜色来区分丢弃策略，考虑了高优先级报文的利益并使其被丢弃的概率相对较小。通过丢弃模板可以配置不同颜色的报文丢弃门限百分比和最大丢弃概率。

4．应用WRED丢弃模板

设备支持在全局、接口、端口队列上应用WRED丢弃模板，可根据需要配置其中一 种或多种。如果在全局（全局应用等效于在所有接口上应用）和接口上同时应用了WRED模板，以接口上应用的模板为准；如果同时在接口、端口队列应用了WRED丢弃模板，系统按照先端口队列后接口的顺序依次匹配报文流，然后依次对匹配WRED丢弃模板的报文流进行拥塞避免控制。

以上4项配置任务的具体配置步骤如表11-13所示。

表11-13 WRED方法拥塞避免配置步骤

配置任务	步骤	命令	说明
（可选）配置端口队列长度	1	**system-view** 例如：<HUAWEI> **system-view**	进入系统视图
	2	**interface** *interface-type interface-number* 例如：[HUAWEI] **interface** gigabitethernet 0/0/1	键入要配置 WRED 方法拥塞避免功能的接口，进入接口视图
	3	**shutdown** 例如：[HUAWEI-GigabitEthernet0/0/1] **shutdown**	关闭以上接口
	4	**qos queue** *queue-index* **length** *length-value* 例如：[HUAWEI-GigabitEthernet0/0/1] **qos queue** 1 **length** 20000	配置接口优先级队列的长度。命令中的参数说明如下。 （1）*queue-index*：指定队列的索引号 （2）*length-value*：指定队列的长度，取值范围为 0～1 000 000 000 的整数，单位是 bytel 缺省情况下，系统统一自动管理各接口的优先级队列长度，可用 **undo qos queue** *queue-index* **length** 命令取消指定接口队列上的优先级队列长度配置
	5	**undo shutdown** 例如：[HUAWEI-GigabitEthernet0/0/1] **undo shutdown**	打开接口
（可选）配置 CFI 作为内部丢弃优先级	6	**dei enable** 例如：[HUAWEI-GigabitEthernet0/0/1] **dei enable**	配置 CFI 作为内部丢弃优先级。缺省情况下，VLAN 中 DEI 字段映射为丢弃优先级的功能未使能，可用 **undo dei enable** 命令去使能将 VLAN 中的 DEI 字段映射为丢弃优先级功能
	7	**quit** 例如：[HUAWEI-GigabitEthernet0/0/1] **quit**	退出接口视图，返回系统视图
配置 WRED 丢弃模板	8	**drop-profile** *drop-profile-name* 例如：[HUAWEI] **drop-profile drop1**	创建 WRED 丢弃模板，并进入丢弃模板视图。参数用来指定 WRED 丢弃模板名称，为 1～31 个字符，不支持空格，不区分大小写 缺省情况下，系统存在一个名为 default 的 WRED 丢弃模板，且只能修改其参数，不能删除，可用 **undo drop-profile** *drop-profile-name* 命令删除指定的 WRED 丢弃模板

（续表）

配置任务	步骤	命令	说明		
配置 WRED 丢弃模板	9	**color** { **green** | **non-tcp** | **red** | **yellow** } **low-limit** *low-limit-percentage* **high-limit** *high-limit-percentage* **discard-percentage** *discard-percentage* 例如：[HUAWEI-drop-drop1] **color green low-limit** 80 **high-limit** 100 **discard-percentage** 10	配置 WRED 丢弃模板的参数，包括丢弃门限百分比和最大丢弃概率。命令中的参数和选项说明如下。 （1）**green**：多选一选项，指定针对绿色报文配置 WRED 参数 （2）**non-tcp**：多选一选项，指定针对非 TCP 报文配置 WRED 参数 （3）**red**：多选一选项，指定针对红色报文配置 WRED 参数 （4）**yellow**：多选一选项，指定针对黄色报文配置 WRED 参数 （5）*owl-limit-percentage*：指定 WRED 丢弃的下限百分比，即当队列中的报文长度占队列长度达到此百分比时，开始进行 WRED 丢弃，取值范围为 0～100 的整数，缺省值为 100 （6）*high-limit-percentage*：指定 WRED 丢弃的上限百分比，即当队列中的报文长度占队列长度达到此百分比时，开始丢弃所有新收到的报文，取值范围为参数 *low-limit-percentage*～100 的整数，缺省值为 100 （7）*discard-percentage*：指定 WRED 的最大丢弃概率，取值范围为 1～100 的整数，缺省值为 100 缺省情况下，WRED 丢弃模板的高低门限百分比以及最大丢弃概率的取值均为 100，可用 **undo color** { **green** | **non-tcp** | **red** | **yellow** }命令恢复对应类型报文的 WRED 丢弃模板参数为缺省值		
应用 WRED 丢弃模板	10	**quit** 例如：[HUAWEI-drop-drop1] **quit**	退出 WRED 丢弃模板视图，返回系统视图	在全局应用 WRED 丢弃模板	可选择其中一种或多种应用方式一
		qos queue *queue-index* **wred** *drop-profile-name* 例如：[HUAWEI]**qos queue** 1 **wred** drop1	将 WRED 丢弃模板应用于交换机上所有端口的指定队列上。命令中的参数说明如下。 （1）*queue-index*：指定要应用 WRED 丢弃模板的所有端口的队列索引号，取值范围为 0～7 的整数 （2）*drop-profile-name*：指定要在所有端口的指定队列上应用的 WRED 丢弃模板的名称，为 1～31 个字符，区分大小写 缺省情况下，端口队列上没有应用 WRED 丢弃模板，可使用 **undo qos queue** *queue-index* **wred** 命令删除在全局端口队列应用的 WRED 丢弃模板		

（续表）

配置任务	步骤	命令	说明		
应用 WRED 丢弃模板	10	**interface** *interface-type interface-number* 例如：[HUAWEI] **interface** gigabitethernet 1/0/1	键入要应用 WRED 丢弃模板的具体接口，进入接口视图	在接口应用 WRED 丢弃模板	可选择其中一种或多种应用方式一
		qos wred *drop-profile-name* 例如：[HUAWEI-GigabitEthernet1/0/1] **qos wred** drop1	将 WRED 丢弃模板应用于接口上所有队列		
		qos queue *queue-index* **wred** *drop-profile-name* 例如：[HUAWEI-GigabitEthernet1/0/1]**qos queue** 1 **wred** drop1	将 WRED 丢弃模板应用于指定端口的指定队列上	在接口队列应用 WRED 丢弃模板	

说明

S5710EI 不支持配置接口优先级队列的长度。另外，以下单板不允许修改队列长 度，缺省支持最大队列长度。

（1）S7700系列交换机的ES1D2G48SBC0单板、ES1D2G48TBC0单板、ES2D2X08SED4单板、ES2D2X08SED5单板和ES1D2L02QFC0单板。

（2）S9300系列交换机的LE0DG48SBC00单板、LE0DG48TBC00单板、LE2D2X08SED4单板、LE2D2X08SED5单板、LE1D2L02QFC0单板和WAN单板。

（3）S9300E系列交换机的LE0DG48SBC00单板、LE0DG48TBC00单板、LH2D2X08SED4单板、LH2D2L02QFC0单板和WAN单板。

（4）S9700系列交换机的EH1D2G48SBC0单板、EH1D2G48TBC0单板、EH1D2X08SED4单板、EH2D2X08SED5单板、EH1D2L02QFC0单板和WAN单板。

【示例 】配置WRED丢弃模板wred1，其中绿色报文的丢弃下限为80%，丢弃上限为100%，最大丢弃概率为10%；黄色报文的丢弃下限为60%，丢弃上限为80%，最大丢弃概率为20%；红色报文的丢弃下限为40%，丢弃上限为60%，最大丢弃概率为40%。

```
<HUAWEI> system-view
[HUAWEI] drop-profile wred1
[HUAWEI-drop-wred1] color green low-limit 80high-limit 100discard-percentage 10
[HUAWEI-drop-wred1] color yellow low-limit 60 high-limit 80discard-percentage 20
[HUAWEI-drop-wred1] color red low-limit 40high-limit 60discard-percentage 40
```

配置好后，可在任意视图下执行display drop-profile [all |name drop-profile-name] 命令查看 WRED 丢弃模板的配置信息；可执行 display qos configuration interface interface-type interface-number命令（已在本章前面介绍）查看指定接口上所有的QoS配置信息。

11.3.4 配置S2700EI系列交换机的拥塞管理

拥塞管理功能实际上就是一个端口队列调度功能。配置拥塞管理后，当网络中发生拥塞时，设备将按照制定的调度策略决定报文转发时的处理次序，以达到高优先级报文优先被调度的目的。S2700EI系列交换机支持的端口对列调度方式包括WRR、PQ+WRR。

S2700EI系列交换机支持4个端口队列，不同的队列可以采用不同的队列调度算法。当调度模式配置为 WRR 时，用户可为每个队列配置权重，S2700 根据权重轮循调度各队列。队列调度时，先调度PQ队列，多个PQ队列按优先级高低顺序进行调度。PQ队列调度完成后，再对 WRR 队列进行加权轮循调度。如果多个接口需要配置相同的队列调度参数，可通过端口组进行配置，以减少重复配置工作。

S2700EI 系列交换机中的队列调度功能配置方法是在系统视图下使用 qos queue queue-index wrr weight weight命令指定端口队列WRR调度的权值。命令中的参数说明如下。

（1）queue-index：指定要配置队列调度功能的端口队列索引号，取值范围为 0～3的整数。

（2）weight：指定对应队列的权重值，取值范围为0～55，权重值越高，越优先被调度。

缺省情况下，WRR调度方式的队列权重为 1，可用undo qos queue queue-indexwrr命令恢复参与WRR调度的某个或某些队列的WRR权值为缺省值1。如果设置某队列权值为0（只有队列2和队列3的权重可以设置为0），说明该队列以PQ方式调度，此时整体调度方式为PQ+WRR。

【示例 】配置队列1的WRR权值为9。

```
<HUAWEI> system-view
[HUAWEI] qos queue 1wrr weight 9
```

11.3.5 配置其他S系列交换机的拥塞管理

除S2700EI系列交换机外，其他华为S系列交换机均支持8个端口队列，不同的队列可以采用不同的队列调度算法。队列调度时，先调度 PQ 队列，多个 PQ 队列按优先级高低顺序进行调度。PQ队列调度完成后，再对WRR或DRR队列进行加权轮循调度。如果多个接口需要配置相同的队列调度参数，可通过端口组进行配置，以减少重复配置工作。但在配置拥塞管理之前，需在报文的入接口上将报文的优先级映射为**PHB**行为，参见本章11.1.4节。

具体的配置步骤因为不同S系列又有所不同，下面分别予以介绍。配置好后，可在任意视图下执行display qos configuration interface [interface-type interface-number] 命令查看指定接口上所有的QoS配置信息；可执行display qos queue statistics interface interface-type interface-number命令查看接口上基于队列的流量统计信息。这两个命令在本章前面均有介绍，不再赘述。

1．S2700-52P-EI/2700-52P-PWR-EI/2710SI/3700SI/3700EI/5700HI/5710EI/S6700系列

在S2700-52P-EI/2700-52P-PWR-EI/2710SI/3700SI/3700EI/5700HI/5710EI/S6700系列交换机中，拥塞管理的配置步骤比较简单，具体如表11-14所示。

表11-14 S2700-52P-EI/2700-52P-PWR-EI/2710SI/3700SI/3700EI系列交换机的拥塞管理配置步骤

步骤	命令	说明
1	**system-view** 例如：<HUAWEI> **system-view**	进入系统视图

（续表）

步骤	命令	说明	
2	**interface** *interface-type interface-number* 例如：[HUAWEI] **interface** gigabitethernet 0/0/1	键入要配置队列调度功能的接口，进入接口视图	
3	**qos** { **pq** | **wrr** | **drr** } 例如：[HUAWEI-GigabitEthernet0/0/1] **qos wrr**	配置端口队列调度模式为 PQ、WRR 或 DRR。缺省情况下，端口队列的调度模式为 WRR 调度模式 【注意】S6700 系列交换机规定只有 0～5 号队列可配置为 WRR 或 DRR 调度，6～7 号队列固定采用 PQ 调度	
4	**qos queue** *queue-index* **wrr weight** *weight* 例如：[HUAWEI-GigabitEthernet0/0/1] **qos queue** 1 **wrr weight** 10	指定端口队列 WRR 调度的权值，只有端口队列调度模式为 WRR 或 PQ+WRR 时，才需要使用此步骤配置。命令中的参数说明如下。 （1）*queue-index*：指定队列的索引号，取值范围 0～7 （2）*weight*：指定对应队列的 wrr 权重值，取值范围为 1～127，权重值越高，越优先被调度 缺省情况下，WRR 调度模式的队列权值为 1，可用 **undo qos queue** *queue-index* **wrr** 命令恢复参与 WRR 调度的某个或某些队列的 WRR 权值为缺省值 1 在采用 WRR 调度方式的前提下，如果设置某队列权值为 0，说明该队列以 PQ 方式调度，此时整体调度模式为 PQ+WRR 方式	二选一
	qos queue *queue-index* **drr weight** *weight* 例如：[HUAWEI-GigabitEthernet0/0/1] **qos queue** 1 **drr weight** 10	指定端口队列 DRR 调度的权值，只有端口队列调度模式为 DRR 或 PQ+DRR 时，才需要使用此步骤配置。命令中的参数说明如下。 （1）*queue-index*：指定队列的索引号 （2）*weight*：指定对应队列的 drr 权重值，取值范围为 1～127，权重值越高，越优先被调度 缺省情况下，DRR 调度模式的队列权值为 1，可用 **undo qos queue** *queue-index* **drr** 命令恢复参与 DRR 调度的某个或某些队列的 DRR 权值为缺省值 1 在采用 DRR 调度方式的前提下，如果设置某队列权值为 0，说明该队列以 PQ 方式调度，此时整体调度模式为 PQ+DRR 方式	

2．S5700SI/5700LI/5700S-LI系列

在 S5700SI/5700LI/5700S-LI 系列交换机中的拥塞管理功能配置方法与前面介绍的S2700-52P-EI/2700-52P-PWR-EI/2710SI/3700SI/3700EI/5700HI/5710EI/S6700 系列交换机中的方法差不多，只不过在这里是在系

统视图下全面配置一个调度模板，然后在具体接口视图下应用所配置的调度模板。具体的配置步骤如表11-15所示。

表11-15 S5700SI/5700LI/5700S-LI系列交换机的拥塞管理配置步骤

步骤	命令	说明
1	**system-view** 例如：<HUAWEI> **system-view**	进入系统视图
2	**qos schedule-profile** *profile-name* 例如: [HUAWEI]**qos schedule-profile** p1	创建全局调度模板，并进入调度模板视图。参数指定所创建的全局调度模板的名称，为 1～16 个字符，不区分大小写
3	**qos** { **pq** | **wrr** | **drr** } 例如：[HUAWEI-qos-schedule-profile-p1] **qos wrr**	配置端口队列调度模式为 PQ、WRR 或 DRR。缺省情况下，端口队列的调度模式为 WRR 调度模式

（续表）

步骤	命令	说明	
4	**qos queue** *queue-index* **wrr weight** *weight* 例如：[HUAWEI-qos-schedule-profile-p1**qos queue** 1 **wrr weight** 10	指定端口队列 WRR 调度的权值，只有端口队列调度模式为 WRR 或 PQ+WRR 时，才需要使用此步骤配置。其他说明参见表 11-16 中第 4 步的说明	二选一
	qos queue *queue-index* **drr weight** *weight* 例如：[HUAWEI-qos-schedule-profile-p1] **qos queue** 1 **drr weight** 10	指定端口队列 DRR 调度的权值，只有端口队列调度模式为 DRR 或 PQ+DRR 时，才需要使用此步骤配置。其他说明参见表 11-16 中第 4 步的说明	
5	**interface** *interface-type interface-number* 例如：[HUAWEI] **interface** gigabitethernet 0/0/1	键入要应用调度模板的接口，进入接口视图	
6	**qos schedule-profile** *profile-name* 例如：[HUAWEI-GigabitEthernet0/0/1] **qos schedule-profile** p1	在以上接口上应用前面配置的调度模板	

3．S7700/9300/9300E/9700系列

在 S7700/9300/9300E/9700 系列交换机的拥塞管理功能的配置方法与前面介绍的S2700-52P-EI/2700-52P-PWR-EI/2710SI/3700SI/3700EI/5700HI/5710EI/S6700 系列交换机中的方法也差不多，只是在这里要区分WAN单板和其他单板，具体配置步骤如表11-16所示。

表11-16 S7700/9300/9300E/9700系列交换机的拥塞管理配置步骤

<table>
<tr><th>步骤</th><th>命令</th><th colspan="2">说明</th></tr>
<tr><td>1</td><td>system-view
例如：<HUAWEI> system-view</td><td colspan="2">进入系统视图</td></tr>
<tr><td>2</td><td>interface interface-type interface-number
例如：[HUAWEI] interface gigabitethernet 0/0/1</td><td colspan="2">键入要配置队列调度功能的接口，进入接口视图</td></tr>
<tr><td>3</td><td>qos queue queue-index wfq weight weight
例如: [HUAWEI-GigabitEthernet0/0/1] qos queue 1 wfq weight 10</td><td>在 WAN 单板中配置端口队列 WFQ 调度的权值。命令中的参数说明如下。
（1）queue-index：指定配置队列调度的隐表索引号，但取值范围仅为 0～4（5～7 号队列固定采用 PQ 调度方式）
（2）weight：指定对应队列的 wrq 权重值，取值范围为 1～100，权重值越高，越优先被调度</td><td rowspan="2">二选一</td></tr>
<tr><td>4</td><td>qos { pq | wrr | drr }
例如：[HUAWEI-GigabitEthernet0/0/1] qos wrr
或
qos { pq { start-queue-index [to end-queue-index] } &<1-8> | { wrr | drr } { start-queue-index [to end-queue-index] } &<1-8> }*
例如：[HUAWEI-GigabitEthernet0/0/1] qos wrr 0 to 5</td><td>在其他单板上配置端口队列调度模式为 PQ、WRR 或 DRR，或者为 PQ+WRR 或 PQ+DRR。两命令中的参数说明如下。
（1）pq：多选一选项，指定采用 PQ 调度模式
（2）wrr：多选一选项，指定采用 WRR 调度模式
（3）drr：多选一选项，指定采用 DRR 调度模式
（4）start-queue-index [to end-queue-index]：指定队列的索引号。其中 start-queue-index 表示第一个队列的索引，end-queue-index 表示最后一个队列的索引</td></tr>
</table>

（续表）

<table>
<tr><th>步骤</th><th>命令</th><th colspan="2">说明</th></tr>
<tr><td>4</td><td>qos { pq | wrr | drr }
例如：[HUAWEI-GigabitEthernet0/0/1] qos wrr
或
qos { pq { start-queue-index [to end-queue-index] } &<1-8> | { wrr | drr } { start-queue-index [to end-queue-index] } &<1-8> }*
例如：[HUAWEI-GigabitEthernet0/0/1] qos wrr 0 to 5</td><td>缺省情况下，端口队列的调度模式为 PQ 调度模式，可用 undo qos { pq | wrr | drr }命令恢复接口下端口队列的调度模式为缺省值
【注意】S7700 系列交换机的 ES1D2X40SFC0 单板、ES1D2L02QFC0 单板，S9700 系列交换机的 EH1D2X40SFC0 单板和 EH1D2L02QFC0 单板，S9300 系列交换机的 LE0DX40SFC00 单板、LE1D2L02QFC0 单板，以及 S9300E 系列交换机的 LE0DX40SFC00 单板、LH2D2L02QFC0 单板的 6～7 号队列固定采用 PQ 调度，只有 0～5 号队列可配置为 WRR 或 DRR 调度</td><td rowspan="3">二选一</td></tr>
<tr><td rowspan="2">5</td><td>qos queue queue-index wrr weight weight
例如：[HUAWEI-qos-schedule-profile-p1qos queue 1 wrr weight 10</td><td>指定端口队列 WRR 调度的权值，只有端口队列调度模式为 WRR 或 PQ+WRR 时，才需要使用此步骤配置。其他说明参见表 11-14 中第 4 步的说明</td></tr>
<tr><td>qos queue queue-index drr weight weight
例如：[HUAWEI-qos-schedule-profile-p1] qos queue 1 drr weight 10</td><td>指定端口队列 DRR 调度的权值，只有端口队列调度模式为 DRR 或 PQ+DRR 时，才需要使用此步骤配置。其他说明参见表 11-14 中第 4 步的说明</td></tr>
</table>

11.3.6 拥塞避免和拥塞管理综合配置示例（一）

本示例拓扑结构如图11-5所示（适用于S5700EI系列），Switch通过接口GE0/0/3与路由器互连，来自Internet的业务有话音、视频、数据，携带的802.1p优先级分别为7、5、2，这些业务可经由路由器和Switch到达用户。现为了减轻网络拥塞造成的影响，保证用户对于高优先级、低延迟业务的服务要求，要求按表11-17 所示配置拥塞避免功能参数，按表11-18所示配置拥塞管理功能参数。

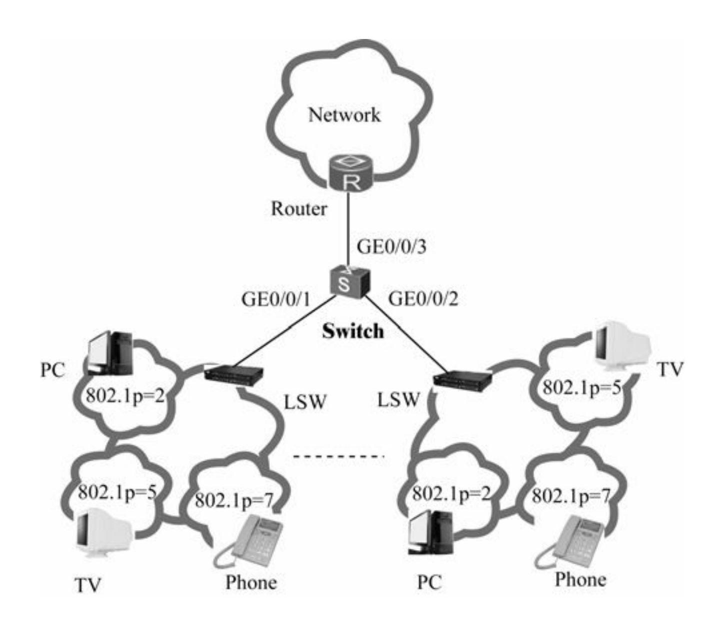


图11-5 拥塞避免和拥塞管理综合配置示例一拓扑结构

表11-17 拥塞避免配置参数

业务类型	颜色	阈值下限	丢弃概率	对应丢弃概率值
话音	黄	1 000	0.78 125%	4
	红	500	6.25%	1
视频	黄	1 000	0.78 125%	4
	红	500	6.25%	1
数据	黄	1 000	0.78 125%	4
	红	500	6.25%	1

表11-18 拥塞管理配置参数

业务类型	服务等级	Wrr 权重
话音	CS7	0
视频	EF	20
数据	AF2	10

1．基本配置思路分析

本示例的基本配置思路如下（不包括VLAN方面的配置）。

（1）首先在Switch设备与路由器连接，Internet流量的入端口GE0/0/3上配置信任报文的802.1p优先级，然后配置基于流的流量监管，对报文进行着色。

（2）全局配置不同 802.1p 优先级报文的拥塞避免功能，即配置各队列的 SRED （S5700EI系列交换机仅支持SRED拥塞避免方法）阈值和丢弃概率。此时只需要配置2、5、7三个队列的调度参数，因为这里有配置优先级与队列的映射，所以采用缺省映射配置，即报文的802.1p优先级与队列号是一一对应的。

（3）在Switch与下级两交换机连接的两个出接口上配置2、5、7三个队列的调度参数。

2．具体配置步骤

（1）配置GE0/0/3入接口信任报文的802.1p优先级。然后按照本章11.3.2节介绍的S5700EI系列交换机基于流的流量监管方法对报文进行着色。

```
<HUAWEI> system-view
[HUAWEI] sysname Switch
```

[Switch] interface gigabitethernet 0/0/3

[Switch-GigabitEthernet0/0/3] trust 8021p

[Switch-GigabitEthernet0/0/3] quit

（2）配置拥塞避免功能，即按照表 10-17 配置 2、5、7队列的 SRED阈值和丢弃概率。

[Switch] qos sred queue 2 red 500 discard-probability 1 yellow 1000 discard-probability 4

[Switch] qos sred queue 5 red 500 discard-probability 1 yellow 1000 discard-probability 4

[Switch] qos sred queue 7 red 500 discard-probability 1 yellow 1000 discard-probability 4

（3）配置拥塞管理功能，即按照表10-18在Switch的GE0/0/1、GE0/0/2出接口上配置各服务等级队列的调度模式。

[Switch] interfacegigabitethernet 0/0/1

[Switch-GigabitEthernet0/0/1] qos wrr

[Switch-GigabitEthernet0/0/1] qos queue7 wrr weight 0

[Switch-GigabitEthernet0/0/1] qos queue5 wrr weight 20

[Switch-GigabitEthernet0/0/1] qos queue2 wrr weight 10

[Switch-GigabitEthernet0/0/1] quit

[Switch] interfacegigabitethernet 0/0/2

[Switch-GigabitEthernet0/0/2] qos wrr

[Switch-GigabitEthernet0/0/2] qos queue7 wrr weight 0

[Switch-GigabitEthernet0/0/2] qos queue5 wrr weight 20

[Switch-GigabitEthernet0/0/2] qos queue2 wrr weight 10

[Switch-GigabitEthernet0/0/2] quit

可通过任意视图命令查看接口出方向队列的全局SRED配置，验证配置结果。具体如下，从中可以看出输出的配置信息与上述配置是一致的，表明配置是正确的。

[Switch] display qos sred

Current sred configuration:

qos sred queue-index 2 red 500 discard-probability 1 yellow 1000 discard-probability 4

qos sred queue-index 5 red 500 discard-probability 1 yellow 1000 discard-probability 4

qos sred queue-index 7 red 500 discard-probability 1 yellow 1000 discard-probability 4

11.3.7 拥塞避免和拥塞管理综合配置示例（二）

本示例拓扑结构如图11-6所示（适用于S5700HI/5710EI/6700/7700/9300/9300E/9700系列交换机）。Switch通过接口GE0/0/3与Router互连，来自Internet的业务有话音、视频、数据，携带的802.1p优先级分别为6、5、2，这些业务可经由Router和Switch到达用户。由于Switch入接口GE0/0/3的速率大于出接口GE0/0/1、GE0/0/2的速率，在这两个出接口处可能会发生拥塞。为了减轻网络拥塞造成的影响，保证用户对于高优先级、低延迟业务的服务要求，现同时配置拥塞避免和拥塞管理功能，配置参数分别如表11-19和表11-20所示。

表11-19 拥塞避免配置参数

业务类型	颜色	阈值下限（%）	阈值上限（%）	丢弃概率
话音	绿	80	100	10
视频	黄	60	80	20
数据	红	40	60	40

表11-20 拥塞管理配置参数

业务类型	服务等级	DRR
话音	EF	0
视频	AF3	100
数据	AF1	50

1．基本配置思路分析

（1）因为在S5700HI/5710EI/6700/7700/9300/9300E/9700系列交换机中支持PHB行为着色，所以需要在Switch上创建并配置DiffServ域，将802.1p优先级映射为PHB行为并着色，并在Switch入接口上绑定DiffServ域。参见本章11.1.4节。

（2）在Switch上配置WRED模板，并在出接口应用WRED模板。

（3）在Switch出接口上配置各服务等级队列的调度参数。

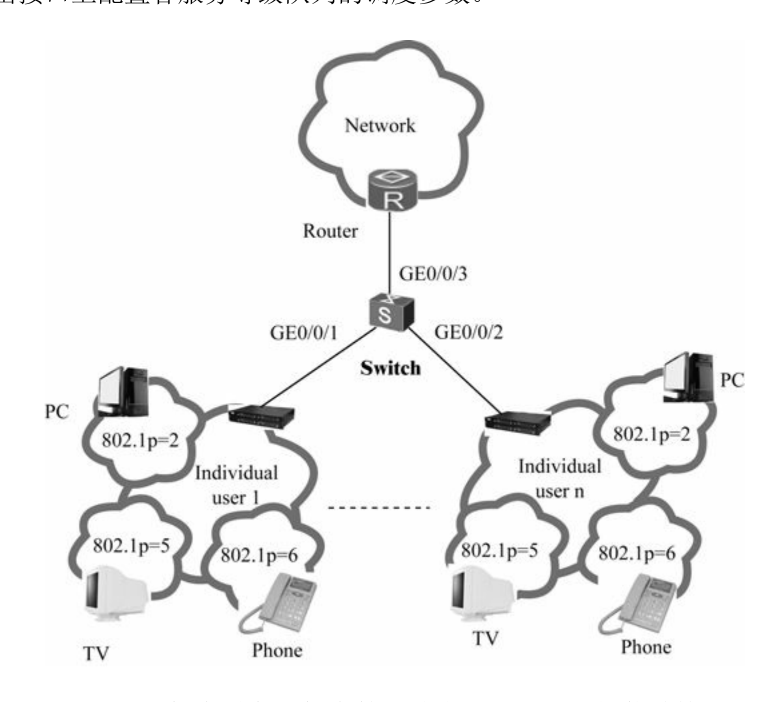


图11-6 拥塞避免和拥塞管理综合配置示例二拓扑结构

2．具体配置步骤

（1）配置基于简单流分类的优先级映射，即创建DiffServ域ds1，将802.1p优先级6、5、2分别映射为PHB行为EF、AF3、AF1，并分别将颜色标记为绿色、黄色、红色。

```
<HUAWEI> system-view
[HUAWEI] sysname Switch
[Switch] diffserv domain ds1
[Switch-dsdomain-ds1] 8021p-inbound 6 phb ef green
[Switch-dsdomain-ds1] 8021p-inbound 5 phb af3 yellow
```

[Switch-dsdomain-ds1] 8021p-inbound 2 phb af1 red

[Switch-dsdomain-ds1] quit

（2）在Switch入接口GE0/0/3上绑定DiffServ域。

[Switch] interfacegigabitethernet 0/0/3

[Switch-GigabitEthernet0/0/3] trust upstreamds1

[Switch-GigabitEthernet0/0/3] trust 8021p inner

[Switch-GigabitEthernet0/0/3] quit

（3）配置拥塞避免，即在Switch上创建WRED模板wred1，并配置wred1的三色报文参数。

[Switch] drop-profile wred1

[Switch-drop-wred1] color green low-limit 80 high-limit 100 discard-percentage10

[Switch-drop-wred1] color yellow low-limit60high-limit 80 discard-percentage20

[Switch-drop-wred1] color red low-limit40 high-limit 60 discard-percentage40

[Switch-drop-wred1] quit

（4）在Switch的GE0/0/1、GE0/0/2出接口上应用WRED模板wred1。

[Switch] interfacegigabitethernet 0/0/1

[Switch-GigabitEthernet0/0/1] qos wredwred1

[Switch-GigabitEthernet0/0/1] qos queue 5 wredwred1

[Switch-GigabitEthernet0/0/1] qos queue 3 wredwred1

[Switch-GigabitEthernet0/0/1] qos queue 1 wredwred1

[Switch-GigabitEthernet0/0/1] quit

[Switch] interfacegigabitethernet 0/0/2

[Switch-GigabitEthernet0/0/2] qos wredwred1

[Switch-GigabitEthernet0/0/2] qos queue 5 wredwred1

[Switch-GigabitEthernet0/0/2] qos queue 3 wredwred1

[Switch-GigabitEthernet0/0/2] qos queue 1 wredwred1

[Switch-GigabitEthernet0/0/2] quit

（5）配置拥塞管理，在 Switch 的 GE0/0/1、GE0/0/2 接口上配置各服务等级队列的调度参数。

[Switch] interfacegigabitethernet 0/0/1

[Switch-GigabitEthernet0/0/1] qos drr

[Switch-GigabitEthernet0/0/1] qos queue 5 drr weight 0 #如果是S7700/9300/9300E/9700系列，本命令要通过以下两条命令来实现：c[Switch-GigabitEthernet0/0/1] qos pq 5，②[Switch-GigabitEthernet0/0/1] qos drr 0 to 4，下同

[Switch-GigabitEthernet0/0/1] qos queue 3 drr weight 100

[Switch-GigabitEthernet0/0/1] qos queue 1 drr weight 50

[Switch-GigabitEthernet0/0/1] quit

[Switch] interfacegigabitethernet 0/0/2

[Switch-GigabitEthernet0/0/2] qos drr

[Switch-GigabitEthernet0/0/2] qos queue 5 drr weight 0

[Switch-GigabitEthernet0/0/2] qos queue 3 drr weight 100

[Switch-GigabitEthernet0/0/2] qos queue 1 drr weight 50

[Switch-GigabitEthernet0/0/2] quit

配置好后，可以通过display diffserv domain name ds1命令查看DiffServ域 ds1的配置信息，验证配置结果。

```
[Switch] display diffserv domain nameds1
diffserv domain name:ds1
8021p-inbound 0 phb be green
8021p-inbound 1 phb af1 green
8021p-inbound 2 phb af1 red
8021p-inbound 3 phb af3 green
8021p-inbound 4 phb af4 green
8021p-inbound 5 phb af3 yellow
8021p-inbound 6 phb ef green
8021p-inbound 7 phb cs7 green
8021p-outbound be green map 0
......
```

同样可通过display drop-profile namewred1命令查看WRED模板配置信息，验证配置结果。

```
[Switch] display drop-profile namewred1
Drop-profile[3] : wred1
Color    Low-limit   High-limit   Discard-percentage
- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
Green       80          100            10
Yellow      60           80            20
Red         40           6             40
Non-tcp     100         100            100
-----------------------------------------------------------------
```

11.4 复杂流策略配置与管理

流策略是指通过将用户流量分类，把具有某类共同特征的报文划分为一类，为相同 类型的流量提供同等的QoS服务，从而针对不同的业务类型提供差分服务。在本书第9章介绍了基于ACL的简化流策略配置与应用方法，本节要介绍复杂流策略（也就是真正的QoS策略）的配置与管理方法。

复杂流策略包含 3 个要素（有关 QoS 流策略的基础知识具体参见第 10 章10.5节）：

（1）流分类

流分类（traffic classifier）用来定义一组流量匹配规则，以对报文进行分类。流分类中各规则之间的关系分为and或or，缺省情况下的关系为and。当流分类中有ACL规则时，报文必须匹配其中一条ACL规则以及所有非ACL规则才属于and类；当流分类中没有ACL规则时，报文必须匹配所有非ACL规则才属于and类；当报文只要匹配了流分类中的一个规则，设备就认为报文属于or类。

（2）流行为

流行为（traffic behavior）用来定义针对某类报文所做的QoS行为。进行流分类是为了有区别地提供服务，它必须与某种流量控制或资源分配的行为关联起来才有意义。

（3）流策略

流策略（traffic policy）用来将指定的流分类和流行为绑定，对分类后的报文执行对应流行为中定义的行为。

11.4.1 配置流分类

配置流分类可以将符合一定规则的报文分为一类，区分出用户流量，是实现差分服务的前提和基础。如果使用 ACL 作为流分类规则，则在配置流分类之前要配置相应的ACL。各个流分类各规则之间属于并列关系，只要匹配规则不冲突，都可以在同一流分类中配置。用户使用时，请根据需要进行配置。

流分类的配置方法很简单，只需以下两步。

（1）先在系统视图下使用 traffic classifier classifier-name [operator {and | or }] 命令创建一个流分类，进入流分类视图。命令中的参数和选项说明如下。

① classifier-name：用来指定所创建的流分类的名称，为 1～31 个字符，不支持空格，区分大小写 。

②and：二选一选项，表示在配置的各流分类中各规则之间关系为“逻辑与”，这样当流分类中有ACL规则时，报文必须匹配其中一条ACL规则以及所有非ACL规则才属于该类；当流分类中没有 ACL规则时，则报文必须匹配所有非 ACL规则才属于该类。

③ or：二选一选项，则表示流分类各规则之间关系是“逻辑或”，即报文只需匹配流分类中的一个或多个规则即属于该类。缺省情况下，流分类中各规则之间的关系为“逻辑与”。

（2）根据实际情况在表11-21中选择配置流分类中的匹配规则。

表11-21 流分类中可以配置的流分类规则

匹配规则	命令	说明
外层 VLAN ID 或基于 QinQ 报文内外两层 Tag 的 VLAN ID	**if-match vlan-id** *start-vlan-id* [**to** *end-vlan-id*] [**cvlan-id** *cvlan-id*] 例如：[HUAWEI-classifier-class1] **if-match vlan-id** 1 **to** 10	在流分类中创建基于 VLAN ID 进行分类的匹配规则。命令中的参数说明如下。 （1）*start-vlan-id* [**to** *end-vlan-id*]：指定匹配报文的外层 VLAN ID。其中 *start-vlan-id* 参数表示起始外层 VLAN ID，*end-vlan-id* 参数表示结束外层 VLAN ID，取值范围均为 1～4 094 的整数。*end-vlan-id* 参数取值必须大于 *start-vlan-id*。若不指定 **to** *end-vlan-id* 可选参数，则表示基于 *start-vlan-id* 参数所指定的外层 VLAN ID 进行流分类 （2）*cvlan-id*：可选参数，指定内层 VLAN ID，取值范围也为 1～4 094 的整数。**S5700SI/5700LI/5700S-LI 系列交换机不支持此参数** 缺省情况下，流分类中没有基于 VLAN ID 进行分类的匹配规则，可用 **undo if-match vlan-id** *start-vlan-id* [**to** *end-vlan-id*] [**cvlan-id** *cvlan-id*] 命令在流分类中删除基于指定 VLAN ID 进行分类的匹配规则
QinQ 报文内外层 VLAN ID	**if-match cvlan-id** *start-vlan-id* [**to** *end-vlan-id*] [**vlan-id** *vlan-id*] 例如：[HUAWEI-classifier-class1] **if-match cvlan-id** 2 **to** 5	配置基于 VLAN ID 进行流分类的匹配规则，**但在 S2700/3700 系列中仅 S2700EI/2710SI/3700EI 系列支持本命令，在 S5700 系列交换机中，S5700SI/5700LI/5700S-LI 系列交换机不支持本命令**。命令中的参数说明如下。 （1）*start-vlan-id* [**to** *end-vlan-id*]：指定匹配 QinQ 报文的内层 VLAN ID。其中 *start-vlan-id* 参数表示起始内层 VLAN ID 取值范围为 1～4 094 的整数，*end-vlan-id* 参数表示结束内层 VLAN ID，取值范围也为 1～4 094 的整数。*end-vlan-id* 参数取值必须大于 *start-vlan-id*。若不指定 **to** *end-vlan-id* 可选参数，则表示基于 *start-vlan-id* 参数所指定的内层 VLAN ID 进行流分类 （2）*vlan-id*：可选参数，指定匹配 QinQ 报文的外层 VLAN ID，取值范围也为 1～4 094 的整数。如果不指定此可选参数，则只匹配 QinQ 报文的内层 VLAN ID 缺省情况下，流分类中没有基于 QinQ 报文内外两层 VLAN ID 进行分类的匹配规则，可用 **undo if-match cvlan-id** *start-cvlan-id* [**to** *end-cvlan-id*] [**vlan-id** *vlan-id*] 命令在流分类中删除基于 QinQ 报文内外两层 VLAN ID 进行分类的匹配规则
VLAN 报文 802.1p 优先级	**if-match 8021p** { *8021p-value* } &<1-8> 例如：[HUAWEI-classifier-class1] **if-match 8021p** 1	配置基于 VLAN 报文的 802.1p 优先级进行流分类的匹配规则，但 **S2700SI 交换机不支持本命令**。命令中的参数说明如下： （1）*8021p-value*：指定配置 VLAN 报文的 802.1p 优先级值，取值范围为 0～7 的整数，值越大优先级越高 （2）&<1-8>：表示可以最多配置 8 个 *8021p-value* 参数值 缺省情况下，流分类中没有基于 VLAN 报文的 802.1p 优先级进行分类的匹配规则，可用 **undo if-match 8021p** 命令在流分类中删除基于 VLAN 报文的 802.1p 优先级进行分类的匹配规则

（续表）

匹配规则	命令	说明
QinQ 报文内层 VLAN 的 802.1p 优先级	**if-match cvlan-8021p** { *8021p-value* } &<1-8> 例如：[HUAWEI-classifier-class1] **if-match cvlan-8021p** 1	配置基于 QinQ 报文内层 802.1p 优先级进行分类的匹配规则，但 **S5700SI/5700LI/5700S-LI 系列交换机不支持本命令**。命令中的两个参数同上一命令说明 缺省情况下，流分类中没有基于 QinQ 报文内层 802.1p 优先级进行分类的匹配规则，可用 **undo if-match cvlan-8021p** 命令在流分类中删除基于 QinQ 报文内层 802.1p 优先级进行分类的匹配规则
丢弃报文	**if-match discard** 例如：[HUAWEI-classifier-class1]**if-match discard**	在流分类中创建基于丢弃报文进行分类的匹配规则。**但在 S2700/3700 系列中仅 S2700EI/2710SI/3700EI 系列支持本命令，在 S5700 系列交换机中，S5700SI/5700LI/5700S-LI 系列交换机不支持本命令**。缺省情况下，流分类中没有基于丢弃报文进行分类的匹配规则，可用 **undo if-match discard** 命令在流分类中删除基于丢弃报文进行分类的匹配规则
QinQ 报文双层 Tag	**if-match double-tag** 例如：[HUAWEI-classifier-class1]**if-match double-tag**	配置基于双层 Tag 进行流分类的匹配规则。**但在 S2700/3700 系列中仅 S2700EI/2710SI/3700EI 系列支持本命令，在 S5700 系列交换机中，S5700SI/5700LI/5700S-LI 系列交换机不支持本命令** 缺省情况下，流分类中没有基于双层 Tag 进行分类的匹配规则，可用 **undo if-match double-tag** 命令在流分类中删除该匹配规则
目的 MAC 地址	**if-match destination-mac** *mac-address* [*mac-address-mask*] 例如：[HUAWEI-classifier-class1]**if-match destination-mac** 0050-b007-bed3 00ff-f00f-ffff	配置基于报文目的 MAC 地址进行流分类的匹配规则。命令中的参数说明如下。 （1）*mac-address*：指定要匹配的目的 MAC 地址，H-H-H 形式，每个 H 代表 4 个十六进制数字 （2）*mac-address-mask*：指定目的 MAC 地址掩码，H-H-H 形式，每个 H 代表 4 个十六进制数字，不能为 0-0-0。MAC 地址的掩码作用与 IP 地址的掩码类似，1 表示匹配该位，0 表示不匹配，可用于确定一组 MAC 地址。用户可以借助 MAC 地址的掩码，实现对目的 MAC 地址中某几位进行精确匹配，具体使用时可在目的 MAC 地址的掩码中将该几位置 1 缺省情况下，流分类中没有基于目的 MAC 地址进行分类的匹配规则，可用 **undo if-match destination-mac** 命令删除基于目的 MAC 地址进行流分类的匹配规则
源 MAC 地址	**if-match source-mac** *mac-address* [*mac-address-mask*] 例如：[HUAWEI-classifier-class1] **if-match source-mac** 0050-b007-bed3 00ff-f00f-ffff	配置基于报文源 AC 地址进行流分类的匹配规则。命令中的参数说明同上一命令 缺省情况下，流分类中没有基于源 AC 地址进行分类的匹配规则，可用 **undo if-match source-mac** 命令删除基于目的 MAC 地址进行流分类的匹配规则
以太网帧头中协议类型字段	**if-match l2-protocol** { **arp** | **ip** | **mpls** | **rarp** | *protocol-value* } 例如：[HUAWEI-classifier-class1] **if-match l2-protocol ip**	配置基于二层报文封装的协议字段进行流分类的匹配规则，但 **S2700SI 交换机不支持本命令**。命令中的参数和选项说明如下。 （1）**arp**：多选一选项，指定基于 ARP 协议字段进行分类 （2）**ip**：多选一选项，指定基于 IP 协议字段进行分类 （3）**mpls**：多选一选项，指定基于 MPLS 协议字段进行分类 （4）**rarp**：多选一选项，指定基于 RARP 协议字段进行分类

（续表）

匹配规则	命令	说明
以太网帧头中协议类型字段	**if-match l2-protocol** { **arp** | **ip** | **mpls** | **rarp** | *protocol-value* } 例如：[HUAWEI-classifier-class1] **if-match l2-protocol ip**	（5）*protocol-value*：多选一选项，指定基于协议类型值进行分类，用十六进制表示，取值范围是 0x0000～0xFFFF，输入时必须以"0x"开始。ARP 协议的类型值为 0x0806，IP 协议的类型值为 0x0800，MPLS 协议的类型值为 0x8847，RARP 协议的类型值为 0x8035。当键入的值小于 0x0600 的时候匹配的是 LLC（Logical Line Control，逻辑链路控制）协议中的 DSAP（Destination Service Access Point，目的服务访问点）和 SSAP（Source Service Access Point，源服务访问点） **本命令为覆盖式命令**，即在同一流分类视图下多次配置基于二层封装的协议字段进行流分类的匹配规则后，**按最后一次配置生效** 缺省情况下，流分类中没有基于二层封装的协议字段进行分类的匹配规则，可用 **undo if-match l2-protocol** 命令用来在流分类中删除基于二层封装的协议字段进行分类的匹配规则
所有报文	**if-match any** 例如：[HUAWEI-classifier-class1] **if-match any**	在流分类中创建基于所有数据报文进行分类的匹配规则，**但 S2700SI 交换机不支持本命令** 缺省情况下，流分类中没有基于所有数据报文进行分类的匹配规则，可用 **undo if-match any** 命令在流分类中删除基于所有数据报文进行分类的匹配规则
IP 报文的 DSCP 优先级	**if-match dscp** *dscp-value* &<1-8> 例如：[HUAWEI-classifier-class1] **if-match dscp** 10	配置基于报文 DSCP 值的匹配规则，**但在 S2700/3700 系列中仅 S2700EI/2710SI/3700EI 系列支持本命令**。命令中的参数说明如下。 （1）*dscp-value* 用来指定要匹配的 DSCP 值，取值范围为 0～63 的整数，也可以为 DSCP 的服务类型名称，它们对应的取值分别为：af11（10）、af12（12）、af13（14）、af21（18）、af22（20）、af23（22）、af31（26）、af32（28）、af33（30）、af41（34）、af42（36）、af43（38）、cs1（8）、cs2（16）、cs3（24）、cs4（32）、cs5（40）、cs6（48）、cs7（56）、default（0）、ef（46）。缺省情况下值为 0 （2）&<1-8>：表示最多可以带 8 个 *dscp-value* 参数值 缺省情况下，流分类中没有基于报文 DSCP 值进行分类的匹配规则，可用 **undo if-match dscp** 命令在流分类删除基于报文 DSCP 值进行分类的匹配规则
IP 报文的 IP 优先级	**if-match ip-precedence** *ip-precedence-value* &<1-8> 例如：[HUAWEI-classifier-class1] **if-match ip-precedence** 1	配置基于 IP 优先级进行分类的匹配规则。命令中的参数说明如下。 （1）*ip-precedence-value*：用来指定要匹配的 IP 优先级值，取值范围为 0～7 的整数 （2）&<1-8>：表示最多可以带 8 个 *ip-precedence-value* 参数值 不能在一个逻辑关系为"与"的流分类中同时配置 **if-match dscp** 和 **if-match ip-precedence** 缺省情况下，流分类中没有基于 IP 优先级进行分类的匹配规则，可用 **undo if-match ip-precedence** 命令在流分类中删除基于 IP 优先级进行分类的匹配规则
报文三层协议类型	**if-match protocol** { **ip** | **ipv6** } 例如：[HUAWEI-classifier-class1] **if-match protocol ip**	配置基于 IPv4 或者 IPv6 协议进行流分类的匹配规则，**但 S2700SI 交换机不支持本命令** 缺省情况下，流分类中没有基于 IP 优先级进行分类的匹配规则，可用 **undo if-match ip-precedence** 命令在流分类中删除基于 IP 优先级进行分类的匹配规则

（续表）

匹配规则	命令	说明
TCP 报文 SYN Flag	**if-match tcp syn-flag** { *syn-flag-value* | **ack** | **fin** | **psh** | **rst** | **syn** | **urg** } 例如：[HUAWEI-classifier-class1] **if-match tcp syn-flag ack**	配置基于 TCP 报文头中的 SYN 标志字段进行流分类的匹配规则，**但 S2700SI 交换机不支持本命令**。命令中的参数和选项说明如下。 （1）*syn-flag-value*：多选一参数，指定用于匹配的 TCP 报文头中 SYN 标志字段的值，取值范围为 0～63 的整数 （2）**ack**：多选一选项，指定用于匹配的 TCP 报文头中 SYN 标志字段的 ACK 标志位 （3）**fin**：多选一选项，指定用于匹配的 TCP 报文头中 SYN 标志字段的 FIN 标志位 （4）**psh**：多选一选项，指定用于匹配的 TCP 报文头中 SYN 标志字段的 PSH 标志位 （5）**rst**：多选一选项，指定用于匹配的 TCP 报文头中 SYN 标志字段的 RST 标志位 （6）**syn**：多选一选项，指定用于匹配的 TCP 报文头中 SYN 标志字段的 SYN 标志位 （7）**urg**：多选一选项，指定用于匹配的 TCP 报文头中 SYN 标志字段的 URG 标志位 缺省情况下，流分类中没有基于 TCP 报文头中的 SYN 标志字段进行分类的匹配规则，可用 **undo if-match tcp** 命令在流分类中删除基于 TCP 报文头中的 SYN 标志字段进行分类的匹配规则
入接口	**if-match inbound-interface** *interface-type interface-number* 例如：[HUAWEI-classifier-class1]**if-match inbound-interface** gigabitethernet 0/0/1	配置基于入接口对报文进行流分类的匹配规则，**但 S2700SI 交换机不支持本命令**。参数 *interface-type interface-number* 用来指定要匹配的入接口类型和编号 缺省情况下，流分类中没有基于入接口对报文进行分类的匹配规则，可用 **undo if-match inbound-interface** 命令在流分类中删除基于入接口对报文进行分类的匹配规则
出接口	**if-match outbound-interface** *interface-type interface-number* 例如：[HUAWEI-classifier-class1]**if-match inbound-interface** gigabitethernet 0/0/2	配置基于出接口对报文进行流分类的匹配规则，**但在 S2700/3700 系列中仅 S2700EI/2710SI/3700EI 系列支持本命令，在 S5700 系列交换机中，S5700SI/5700LI/5700S-LI 系列交换机不支持本命令**。参数 *interface-type interface-number* 用来指定要匹配的出接口类型和编号 缺省情况下，流分类中没有基于入接口对报文进行分类的匹配规则，可用 **undo if-match outbound-interface** 命令在流分类中删除基于出接口对报文进行分类的匹配规则
ACL 规则	**if-match acl** { *acl-number* | *acl-name* } 例如：[HUAWEI-classifier-class1] **if-match acl** 2001	使用 ACL 作为流分类规则，{ *acl-number* | *acl-name* }参数分别用来指定要匹配的 ACL 号或 ACL 名称 必须先配置相应的 ACL 规则，设备支持：基本 IPv4 ACL、高级 IPv4 ACL、二层 ACL 和用户自定义 ACL，参见本书第 9 章
ACL6 规则	**if-match ipv6 acl** { *acl-number* | *acl-name* } 例如：[HUAWEI-classifier-class1] **if-match ipv6 acl** 3001	使用 ACL6 作为流分类规则，{ *acl-number* | *acl-name* }参数分别用来指定要匹配的 ACL 号或 ACL 名称 必须先配置相应的 ACL6 规则，设备支持基本 IPv6 ACL 和高级 IPv6 ACL，**但 S2700/3700 系列交换机不支持 IPv6 ACL**

【示例 1】定义流分类 c1的匹配规则为匹配VLAN ID为 2的报文。

<HUAWEI> system-view

[HUAWEI] traffic classifier c1 operator and

[HUAWEI-classifier-c1] if-match vlan-id 2

【示例 2】定义流分类 c1的匹配规则为匹配内层VLAN ID为 100的QinQ报文。

<HUAWEI>system-view

[HUAWEI] traffic classifier c1 operator and

[HUAWEI-classifier-c1] if-match cvlan-id 100

【示例 3】定义流分类 c1的匹配规则为匹配内层VLAN ID范围为 100到200且外层VLAN ID为 300的QinQ报文。

<HUAWEI>system-view

[HUAWEI] traffic classifier c1 operator and

[HUAWEI-classifier-c1] if-match cvlan-id 100 to 200 vlan-id 300

【示例 4】定义流分类c1的匹配规则为匹配802.1p优先级值为1的VLAN报文。

<HUAWEI>system-view

[HUAWEI] traffic classifier c1 operator and

[HUAWEI-classifier-c1] if-match 8021p 1

【示例 5】定义流分类c1的匹配规则为匹配目的MAC地址为0050-ba27-bed3的报文。

<HUAWEI>system-view

[HUAWEI] traffic classifier c1 operator and

[HUAWEI-classifier-c1] if-match destination-mac 0050-ba27-bed3

【示例 6】定义流分类c1的匹配规则为匹配目的MAC地址为XX50-bXX7-bed3的报文（X表示为任意十六进制）。

<HUAWEI>system-view

[HUAWEI] traffic classifier c1 operator and

[HUAWEI-classifier-c1] if-match destination-mac 0050-b007-bed3 00ff-f00f-ffff

【示例 7】定义流分类 c1 的匹配规则为匹配二层封装的协议字段为 ARP 的报文。

<HUAWEI>system-view

[HUAWEI] traffic classifier c1 operator and

[HUAWEI-classifier-c1] if-match l2-protocol arp

配置好流分类后，可以通过执行 display traffic classifier user-defined [classifier- name] 任意视图命令查看设备上的流分类信息。

11.4.2 配置流行为

配置流行为即为符合流分类规则的流量指定后续行为，是配置流策略的前提条件。本节将介绍以下流行为的配置方法，可根据实际需要选择配置。在配置流行为时，各行为属于叠加关系，只要不冲突，都可以在同一流行为中配置。

1．报文过滤

配置报文过滤后，设备将对符合流分类规则的报文进行过滤，从而实现对网络流量的控制。

2．重标记

通过配置重标记，设备对符合流分类规则的报文的指定字段进行设置，如VLAN报文的802.1p优先级、IP报文的DSCP和内部优先级。但重标记报文某些字段，不会影响当前设备对报文的**QoS**处理，仅会影响下游设备对报文的**QoS**处理 。

3．重定向

通过配置重定向，设备将符合流分类规则的报文重定向到 CPU、指定的下一跳 地址或指定接口，但包含重定向行为的流策略只能在全局、接口或 **VLAN** 的入方向上应用 。

4．流量监管

流量监管是一种通过对流量规格的监督，来限制流量及其资源使用的流量控制行为。通过配置流量监管，设备对符合流分类规则的报文的流量进行监督，对于超过规格的流量，可以采取丢弃、重标记颜色、重标记服务级别等行为。

5．流镜像

配置流镜像后，设备将符合流分类规则的所有报文镜像到监控端口。

6．流量统计

配置流量统计后，设备将对符合流分类规则的报文进行流量统计，可以帮助用户了解应用流策略后报文通过和被丢弃的情况，由此分析和判断流策略的应用是否合理，也有助于进行相关的故障诊断与排查。

7．禁止MAC地址学习

在网络比较稳定、报文的 MAC 地址相对固定的情况下，设备没有必要继续学习其他所有报文的 MAC 地址。此时通过对流策略下所有流分类禁止 MAC 地址学习功能，既可以节省MAC地址表项开支，又可以提高设备的运行效率。

某些非法用户有时会采用频繁变换 MAC 地址的方式对网络进行攻击，此时通过对流策略下所有流分类禁止 MAC 地址学习功能，可以避免此类攻击所造成的设备 MAC地址表项溢出的问题，保护设备性能不受影响。

以上流行为的配置步骤如表11-22所示。

表11-22 流行为配置步骤

配置任务	步骤	命令	说明
公共配置	1	**system-view** 例如：<HUAWEI> **system-view**	进入系统视图
	2	**traffic behavior** *behavior-name* 例如：[HUAWEI] **traffic behavior** b1	创建一个流行为，进入流行为视图。参数用来指定所创建的流行为的名称，为 1～31 个字符，不支持空格，**区分大小写**
配置报文过滤	3	**permit** | **deny** 例如：[HUAWEI-behavior-b1] **deny**	如果执行 **permit** 命令，则对符合流分类的报文不做任何行为，按原来的策略转发。在流行为中，**permit** 行为和其他流行为一起配置时将依次执行这些行为；如果执行 **deny** 命令，则禁止符合流分类规则的报文通过。**deny** 行为和其他流行为互斥，即使配置其他行为也不会生效（**流量统计和流镜像除外**） 【注意】为匹配 ACL 规则的报文指定报文过滤行为时，如果此 ACL 中的 rule 规则配置为 **permit**，则设备对此报文采取的行为由此处流行为中配置的 **deny** 或 **permit** 决定；如果此 ACL 中的 rule 规则配置为 **deny**，则无论在流行为中配置 **deny** 或 **permit** 行为，此报文都被丢弃

（续表）

配置任务	步骤	命令	说明
配置重标记（可选择配置一个或多个重标记行为）	4	**remark 8021p** [*8021p-value* | **inner-8021p**] 例如：[HUAWEI-behavior-b1] **remark 8021p** 4	将符合流分类的报文重新标记 802.1p 优先级。命令中的参数和选项说明如下。 （1）*8021p-value*：二选一参数，用来指定重新标记的 802.1p 优先级值，取值范围为 0～7 的整数 （2）**inner-8021p**：二选一选项，指定重标记的 802.1p 优先级值是从内层继承的 缺省情况下，流行为中没有重标记 VLAN 报文 802.1p 优先级的行为，可用 **undo remark 8021p** 命令在流行为中删除重标记 VLAN 报文 802.1p 优先级的行为 【注意】**remark 8021p inner-8021p** 命令仅能在入方向应用，包含 **remark 8021p** 行为的流策略应用在接口出方向时，出接口 VLAN 必须工作在带标签方式
		remark dscp { *dscp-name* | *dscp-value* } 例如：[HUAWEI-behavior-b1] **remark dscp af13**	将符合流分类的报文重新标记 DSCP 值，**在 S2700/3700 系列中仅 S2700EI/2710SI/3700EI 系列支持本命令，在 S5700 系列交换机中，S5700SI/5700LI/5700S-LI 系列交换机不支持本命令**。命令中的参数说明如下。 （1）*dscp-name*：二选一参数，指定重标记为对应名称的 DSCP 值，可以是 **ef**、**af11**、**af12**、**af13**、**af21**、**af22**、**af23**、**af31**、**af32**、**af33**、**af41**、**af42**、**af43**、**cs1**、**cs2**、**cs3**、**cs4**、**cs5**、**cs6**、**cs7** 或 **default** （2）*dscp-value*：二选一参数，指定重标记为对应 DSCP 值，取值范围为 0～63 缺省情况下，流行为中没有重标记 IP 报文的 DSCP 优先级的行为，可用 **undo remark dscp** 命令在流行为中删除重标记 IP 报文的 DSCP 优先级的行为
		remark cvlan-id *cvlan-id* 例如：[HUAWEI-behavior-b1] **remark cvlan-id** 10	将符合流分类的 QinQ 报文重新标记内层 VLAN 标签的值，**但 S2700/3700 系列交换机不支持本命令，在 5700EI 系列交换机上，本命令不支持应用到入方向**。命令中的 *cvlan-id* 参数用来指定重标记后的内层 VLAN ID，取值范围为 1～4 094 缺省情况下，流行为中没有重标记 QinQ 报文中的内层 VLAN 标签值的行为，可用 **undo remark cvlan-id** 命令在流行为中删除配置重标记 QinQ 报文中的内层 VLAN 标签值的行为

（续表）

配置任务	步骤	命令	说明
配置重标记（可选择配置一个或多个重标记行为）	4	**除 S2700-52P-EI 和 S2700-52P-PWR-EI 之外的其他 S2700EI 系列交换机，以及 S5700SI/5700LI/5700S-LI 系列交换机：** **remark local-precedence** { *local-precedence-name* | *local-precedence-value* } 例如：[HUAWEI-behavior-b1] **remark local-precedence** 3 **其他 S 系列交换机：** **remark local-precedence** { *local-precedence-name* | *local-precedence-value* } [**green** | **yellow** | **red**] 例如：[HUAWEI-behavior-b1] **remark local-precedence** 3 **green**	将符合流分类的报文重标记内部优先级值。命令中的参数和选项说明如下。 （1）*local-precedence-name*：二选一参数，指示以本地优先级名称重标记本地优先级，取值可为 **af1**、**af2**、**af3**、**af4**、**be**、**cs6**、**cs7** 或 **ef** （2）*local-precedence-value*：二选一参数，指示以本地优先级值重标记本地优先级，取值范围为 0～7 的整数，值越大优先级越高 （3）[**green** | **yellow** | **red**]：可选项，指定所重标记后的内部优先级对应的报文颜色分别为绿色、黄色或红色 缺省情况下，流行为中没有重标记内部优先级的行为，可用 **undo remark local-precedence** 命令在流行为中删除重标记内部优先级的行为
		remark ip-precedence *ip-precedence* 例如：[HUAWEI-behavior-b1] **remark ip-precedence** 3	将符合流分类的报文重标记 IP 优先级值。**仅 S2700-52P-EI/2700-52P-PWR-EI/2710SI/300SI/3700EI/5700/6700 系列交换机支持本命令**。参数用来指定重标记后的 IP 优先级值，取值范围为 0～7 的整数，值越大，优先级越高 缺省情况下，流行为中没有重标记报文的 IP 优先级的行为，可用 **undo remark ip-precedence** 命令取消标记报文的 IP 优先级
		remark destination-mac *mac-address* 例如：[HUAWEI-behavior-b1] **remark destination-mac** 0050-b007-bed3	将符合流分类的报文重标记目的 MAC 地址。除 S2700SI /5700SI/5700LI/5700S-LI 系列外的其他所有 S 系列交换机均支持本命令。参数 *mac-address* 用来指定重标记后的目的 MAC 地址（**必须为单播 MAC 地址**），H-H-H 形式，每个 H 代表 4 个 16 进制数字 缺省情况下，流行为中没有重标记报文目的 MAC 地址的行为，**undo remark destination-mac** 命令在流行为中删除重标记报文目的 MAC 地址的行为
		remark vlan-id *vlan-id* 例如：[HUAWEI-behavior-b1] **remark vlan-id**	将符合流分类的报文重标记 VLAN 标签值。参数用来指定重标记后的 VLAN ID，取值范围为 1～4 094 缺省情况下，流行为中没有重标记 VLAN 报文的 VLAN 标签值的行为，可用 **undo remark vlan-id** 命令在流行为中删除重标记 VLAN 报文的 VLAN 标签值的行为 【注意】只要在流策略中包含 **remark vlan-id** 行为，当应用在接口出方向时，出接口必须工作在带标签方式

（续表）

配置任务	步骤	命令	说明
配置重定向（可选择配置一个或多个重定向行为，对于 S2700 系列交换机，只有 S2700-52P-EI 和 S2700-52P-PWR-EI 系列交换机支持重定向行为）	5	**redirect cpu** 例如：[HUAWEI-behavior-b1] **redirect cpu**	将符合流分类的报文重定向到 CPU，**S3700SI/5700SI/5700LI/5700S-LI 系列交换机也不支持此命令**。应用包含本流行为的流策略后，会将符合流分类规则的报文重定向到 CPU 缺省情况下，流行为中没有将报文重定向到 CPU 的行为，可用 **undo redirect** 命令在流行为中删除重定向配置
		redirect ip-nexthop *ip-address* <1-4> [**forced**] 例如：[HUAWEI-behavior-b1] **redirect ip-nexthop** 10.10.10.1	将符合流分类的报文重定向到下一跳，**S3700SI**/5700SI/5700LI/5700S-LI 系列交换机也不支持此命令。命令中的参数说明如下。 （1）*ip-address*：指定重定向的下一跳 IP 地址 （2）<1-4>：表示最多可以带 4 个 *ip-address* 参数值 （3）**forced**：指定当下一跳不存在的时候，直接丢弃该报文 【说明】当存在多个下一跳时，设备按照主备方式对报文进行重定向转发。一个流行为中最多可以配置 4 个下一跳，设备根据下一跳的配置顺序确定主备链路，先配置的下一跳 IP 地址优先级较高。配置的第一个下一跳 IP 地址作为主用链路，其他链路作为备用链路。当主用链路 Down 之后，则自动选取优先级高的下一跳作为新的主链路 缺省情况下，流行为中没有将报文重定向到下一跳 IP 地址的行为，可用 **undo redirect** 命令在流行为中删除重定向配置。不能在同一流策略中同时配置 **redirect ip-nexthop** 和 **remark destination-mac** 流行为
		redirect ip-multihop { **nexthop** *ip-address* } <2-4> 例如：[HUAWEI-behavior-b1] **redirect ip-multihop nexthop 13.1.42.1 nexthop 23.2.12.3 nexthop** 32.1.1.2	将符合流分类的报文重定向到配置的多个下一跳中的一个，**S3700SI/5700SI/5700LI/5700S-LI 系列交换机也不支持此命令**。命令中的参数说明如下。 （1）**nexthop** *ip-address*：指定重定向的下一跳 IP 地址 （2）<2-4>：表示最少要带 2 个 **nexthop** *ip-address* 参数值，最多可带 4 个 【说明】如果配置了多个下一跳，设备将按照等价路由负载分担方式对报文进行重定向转发，即设备按照报文的源 IP 地址（不管流量大小）并根据 HASH 算法在多个下一跳中选择一个进行转发。源 IP 地址相同的流量，则不管流量多大都是选择同一个下一跳转发。使用重定向到多个下一跳的正常转发过程中，如果当前下一跳对应的出接口状态突然为 Down，或路由突然发生了改变，设备可将链路快速切换到当前可用的某个下一跳对应的出接口上。如果设备上没有命令中配置的下一跳 IP 地址所对应的 ARP 表项，使用此命令能配置成功，但重定向不能生效，设备仍按报文原来的目的地址转发，直到设备上有对应的 ARP 表项

（续表）

配置任务	步骤	命令	说明
配置重定向（可选择配置一个或多个重定向行为，对于S2700系列交换机，只有S2700-52P-EI和S2700-52P-PWR-EI系列交换机支持重定向行为）	5	**redirect ip-multihop** { **nexthop** *ip-address* } &<2-4> 例如：[HUAWEI-behavior-b1] **redirect ip-multihop nexthop** 13.1.42.1 **nexthop** 23.2.12.3 **nexthop** 32.1.1.2	缺省情况下，流行为中没有将报文重定向到多个下一跳 IP 地址的行为，可用 **undo redirect** 命令在流行为中删除重定向配置 不能在同一流策略中同时配置 **redirect ip-multihop** 和**remark destination-mac**流行为
		redirect interface *interface-type interface-number* 例如：[HUAWEI-behavior-b1] **redirect interface** gigabitethernet 0/0/1	将符合流分类的报文重定向到指定接口，**S3700SI /5700SI/5700LI/5700S-LI 系列交换机也不支持此命令**。如果此接口 Down 了，就在此接口丢包，流量不会切换到原转发路径 缺省情况下，流行为中没有将报文重定向到指定接口的行为，可用 **undo redirect** 命令在流行为中删除重定向配置
配置流量监管（根据设备类型选择一个配置，但S2700SI系列不支持）	6	**除 S2700-52P-EI 和 S2700-52P-PWR-EI 之外的其他 S2700EI 系列交换机：** **car** [**aggregation**]**cir** *cir-value* **cbs** *cbs-value* **S2700-52P-EI/2700-52P-PWR-EI/2710SI/3700SI/3700EI/5700EI 系列交换机：** **car** [**aggregation**]**cir** *cir-value*[**pir** *pir-value*] [**cbs** *cbs- value* **pbs** *pbs-value*] [**green** { **discard** |**pass** [**remark-dscp** *dscp-value* | **remark-8021p***8021p-precedence*] }] [**yellow** { **discard** | **pass**[**remark-dscp** *dscp- value* | **remark-8021p** *8021p-precedence*] }] [**red** { **discard** | **pass**[**remark-dscp** *dscp-value* | **remark-8021p** *8021p-precedence*] }] **S5700LI/5700S-LI 系列交换机：** **car cir** *cir-value*[**pir** *pir-value*] [**cbs** *cbs-value* **pbs** *pbs-value*] [**green pass**] [**yellow** { **discard** | **pass** [**remark-dscp** *dscp-value* | **remark-8021p** *8021p-precedence*] }] [**red** { **discard** | **pass**[**remark-dscp** *dscp-value* | **remark-8021p** *8021p-precedence*] }] **S5700SI 系列交换机：** **car** [**aggregation**] **cir** *cir-value* [**pir** *pir-value*] [**cbs** *cbs-value* **pbs** *pbs-value*] [**green pass**] [**yellow** { **discard** | **pass**[**remark-dscp** *dscp-value* |**remark-8021p** *8021p-precedence*] }] [**red** { **discard** | **pass**[**remark-dscp** *dscp-value* | **remark-8021p** *8021p-precedence*] }]	在流行为中创建流量监管行为。各命令中的参数和选项说明如下。 （1）**aggregation**：可选项，指定所配置的 CAR 为聚合 CAR。在聚合 CAR 的情况下，同类型的规则应用到多个接口上，只占用一个 CAR 资源，这些接口的流量都受这个 CAR 约束 （2）*cir-value*：指定承诺信息速率，即保证能够通过的速率，单位是 kbit/s。S5700HI/5700LI/5700S-LI/5710EI/6700 系列的取值范围为：8～10 000 000 的整数；S2700/3700/5700SI/5700EI 系列的取值范围为 64～10 000 000 的整数；S7700/9300/9300/9700 系列的取值范围为 8～4 294 967 295 的整数 （3）*pir-value*：可选参数，指定峰值信息速率，即最大能够通过的速率，单位是 kbit/s。它的取值范围也因为不同系列有所不同，具体同 *cir-value* 参数，但必须大于或等于 *cbs-value*。缺省情况下，*pir-value* 的值等于 *cir-value* 的值 （4）*cbs-value*：可选参数，指定承诺突发尺寸，即瞬间能够通过的承诺突发流量，单位是 byte。S2700 系列的取值范围为 8 192～4 294 967 295 的整数，S3700/5700/6700 系列的取值范围为 4 000～4 294 967 295 的整数，S7700/9300/9300/9700 系列的取值范围为 10 000 ～ 4 294 967 295 的整数。在 S2700/3700/5700/6700 系列中，如果 *cir-value*≤4 000kbit/s，则 *cbs-value* 的缺省值为 4 000byte；如果 *cir-value*>4 000kbit/s，则 *cbs-value* 的缺省值等于 *cir-value* 的值；在 S7700/9300/9300/9700 系列，单令牌桶时，*cbs-value* 缺省为 *cir-value* 的 188 倍；双令牌桶时，*cbs-value* 缺省为 *cir-value* 的 125 倍

（续表）

配置任务	步骤	命令	说明
配置流量监管（根据设备类型选择一个配置，但S2700SI 系列不支持）	6	**S5700HI/5710EI/6700 交换机：** **car cir** *cir-value* [**pir** *pir-value*] [**cbs** *cbs-value* **pbs** *pbs-value*] [**green** {**discard** | **pass**[**remark-dscp** *dscp-value* | **remark-8021p** *8021p-precedence*] }] [**yellow** { **discard** | **pass**[**remark-dscp** *dscp-value* | **remark-8021p** *8021p-precedence*] }] [**red** { **discard** | **pass**[**remark-dscp** *dscp-value* | **remark-8021p** *8021p-precedence*] }] **S7700/9300/9300E/9700 系列交换机：** **car cir** *cir-value*[**pir** *pir-value*] [**cbs** *cbs-value* **pbs** *pbs-value*] [**share**] [**mode** { **color-blind** | **color-aware** }] [**green** { **discard** | **pass**[**service-class** *class* **color** *color*] } | **yellow** { **discard** | **pass** [**service-class** *class* **color** *color*] } | **red** {**discard** | **pass** [**service-class** *class* **color** *color*] }]* 例如：[HUAWEI-behavior-b1]**car cir** 200000 **pir** 2500000 **green pass yellow pass red discard**	（5）*pbs-value*：可选参数，指定峰值突发尺寸，即瞬间能够通过的峰值突发流量，单位是 byte。S2700/3700/5700/6700 系列的取值范围为 4 000～4 294 967 295 的整数，S7700/9300/9300/9700 系列的取值范围为 10 000～4 294 967 295 的整数。如果不配置此可选参数，则 S2700/3700/5700/6700 系列交换机缺省值与配置的 *cir-value* 和 *pir-value* 有关：如果未配置 *pir-value* 参数，且 *cir-value**125≤4 000kbit/s，则 *pbs-value* 的缺省值为 4 000byte；如果 *pir-value**125>4 000kbit/s，则 *pbs-value* 的缺省值等于 *cir-value* 的 125 倍。如果配置了 *pir-value* 参数，且 *pir-value**125≤4 000kbit/s，则 *pbs-value* 的缺省值为 4 000byte；如果 *pir-value**125>4 000kbit/s，则 *pbs-value* 的缺省值等于 *pir-value* 的 125 倍。在 S7700/9300/9300/9700 系列中，如果不配置 *pir-value* 参数，则 *pbs-value* 缺省为 *cir-value* 的 313 倍；如果配置 *pir-value* 参数，则 *pbs-value* 缺省为 *pir-value* 的 125 倍 （6）**green**、**yellow**、**red**：可选项，指定报文的颜色，由本命令中的参数 **cbs** *cbs-value*、**pbs** *pbs-value* 确定。缺省情况下，绿色、黄色报文被允许通过，红色报文被丢弃 （7）**discard** | **pass**：表示丢弃报文或者允许报文通过 （8）**remark-8021p** *8021p-precedence*：二选一参数，指定重标记报文的 802.1p 优先级，取值范围为 0～7 的整数 （9）**remark-dscp** *dscp-value*：二选一参数，指定重标记报文的 DSCP 优先级，取值范围为 0～63 的整数 （10）**service-class** *class*：可选参数，指定服务等级，取值可为 **af1**、**af2**、**af3**、**af4**、**be**、**cs6**、**cs7**、**ef** 八种服务等级 （11）**color** *color*：可选参数，指定由 **service-class** *class* 参数配置的服务等级所对应的颜色，取值包括 **green**、**yellow**、**red** 三种颜色（优先级依次降低），报文最终的颜色是由 DiffServ 域中定义的报文颜色和流量监管中定义的报文颜色共同决定，取优先级低的颜色 缺省情况下，流行为中没有流量监管行为，可用 **undo car** 命令在流行为中删除流量监管行为

（续表）

配置任务	步骤	命令	说明
配置流镜像（S2700SI 系列不支持）	7	**mirroring to observe-port** *observe-port-index* 例如：[HUAWEI-behavior-b1] **mirroring to observe-port** 2	将满足流分类规则的所有流镜像到监控端口。参数用来指定要镜像到的监控端口索引号，取值范围不同系列有所不同：S2700/5700SI/5700S-LI/5700LI/6700 系列仅为 1，S5700HI/5710EI 系列的取值范围是 1～2 的整数，S3700/5700EI 系列的取值范围为 1～4 的整数，S7700/9300/9300E/9700 系列的取值范围为 1～8 的整数 需要先执行 **observe-port**（本地镜像）或 **observe-port**（远程镜像）命令创建监控端口，才能在流行为中配置镜像到该监控端口 缺省情况下，系统未配置将满足规则的流镜像到监控端口，可用 **undo mirroring** 命令取消将满足规则的流镜像到观察端口
配置流量统计	8	**statistic enable** 例如：[HUAWEI-behavior-b1] **statistic enable**	使能流量统计功能，**但 S2700SI 系列不支持**。缺省情况下，流行为中未使能流量统计功能，可用 **undo statistic enable** 命令去使能流策略的统计功能
配置禁止 MAC 地址学习	9	**mac-address learning disable** 例如：[HUAWEI-behavior-b1] **mac-address learning disable**	在流行为中去使能 MAC 地址学习功能，**但 S2700/3700 系列中仅 S2700EI/2710SI/ 3700EI 系列支持，在 S5700 系列交换机中 S5700SI/5700LI/5700S-LI 系列交换机不支持** 缺省情况下，流行为中的 MAC 地址学习功能处于使能状态，**undo mac-address learning disable** 命令使能流行为中的 MAC 地址学习功能
配置封装外层 VLAN 标签	10	**nest top-most vlan-id** *vlan-id* 例如：[HUAWEI-behavior-b1] **nest top-most vlan-id** 3	配置封装外层 VLAN 标签，**仅 S7700/ 9300/9300E/9700 系列支持**，主要用于实现基于流的灵活 QinQ 功能。参数用来指定要封装的外层 VLAN ID，取值范围为 1～4 094 的整数 缺省情况下，流行为中没有配置创建外层 VLAN 标签的行为，可用 **undo nest** 命令在流行为中取消创建外层 VLAN 标签的行为

配置好流行为后可以通过 **display traffic behavioruser-defined** [behavior-name] 任意视图命令查看流行为的配置信息。

11.4.3 配置流策略

通过配置流策略，将流分类和流行为绑定起来，形成完整的策略。这一步的配置很简单，只需先在系统视图下通过 traffic policy traffic-policy-name [match-order { auto | config }] 命令创建流策略，并进入流策略视图，然后在流策略视图下通过classifier classifier-name behavior behavior-name 命令将前面两节配置的流分类和流行为进行关联即可。

traffic policy 命令中的参数和选项说明如下。

（1）traffic-policy-name：指定创建的流策略的名称，为1～31个字符，不支持空格， 区分大小写 。

（2）[match-order {auto | config }] ：可选项，指定流策略规则匹配顺序，如果选择二选一选项auto，则流策略规则匹配顺序是由系统预先指定的流分类优先级决定的，该优先级排序如下：基于二层和三层信息流分类 >基于二层信息流分类 >基于三层信息流分类；如果选择二选一选项config，则流策略规则匹配顺序是由流分类配置的先后顺序决定的。但在**S2700/3700/5700SI/5700EI/5700LI/5700S-LI**系列交换机中不支持此选项 。

缺省情况下，系统未创建任何流策略，可用undo traffic policy policy-name命令删除指定的流策略。但如果流策略已经应用到全局、接口板、接口或VLAN，则一般不允许删除该策略。如果不得不删除，则需要先在相应的视图下执行undo traffic-policy命令取消对该策略的应用，然后到系统视图下执行undo traffic policy 命令完成删除。

【示例 1】创建根据流分类配置顺序进行规则匹配的流策略 p1，并关联已创建的流分类c1和流行为b1。

<HUAWEI> system-view

[HUAWEI] traffic policy p1 match-order config

[HUAWEI-trafficpolicy-p1] classifier c1 behavior b1

【示例 2】删除已经应用在接口GE1/0/1入方向上的流策略p1。

<HUAWEI>system-view

[HUAWEI] interface gigabitethernet 1/0/1

[HUAWEI-GigabitEthernet1/0/1] undo traffic-policy inbound

[HUAWEI-GigabitEthernet1/0/1] quit

[HUAWEI] undo traffic policy p1

11.4.4 应用流策略

绑定了流行为与流分类的完整流策略可应用到交换机全局、接口或VLAN上，实现针对不同业务的差分服务。下面分别介绍不同应用方式的具体配置方法。

说明

在报文同时匹配多个流策略时。

● 如果这些流策略的分类规则属于同一类，即匹配规则同属于自定义 ACL 规则、二层规则或三层规则时，只会有一个流策略生效，生效的优先级与应用的对象有关，生效优先级：接口>VLAN>全局。

● 如果这些流策略的分类规则不属于同一类，对于彼此不冲突的行为，流策略都会生效；对于彼此冲突的行为，流策略生效优先级与规则有关，生效优先级：自定义ACL规则>二层规则+三层规则>二层规则>三层规则。

1．在全局应用流策略

在全局应用流策略是指在交换机所有端口的某个方向上应用所创建的流策略。不同S系列交换机的全局应用流策略配置命令有所不同。

（1）在 S2700/3700 系列交换机中（**S2700SI** 交换机不支持流策略应用 ）。在系统视图下使用**traffic-policy** policy-name **global inbound** 命令在交换机上所有端口入方向应用流策略。缺省情况下，没有应用任何流策略，可用**undo traffic-policy global inbound** 命令删除在全局应用的流策略。

（2）在 S5700/6700/7700/9300/9300E/9700 系列交换机中。在系统视图下使用 traffic-policypolicy-nameglobal { inbound | outbound } [slot slot-id] 命令交换机或某个单板上所有端口入方向或出方向应用流策略。

说明

全局或单板的每个方向上只能应用一个流策略，如果在全局某方向应用了流策略， 则不能在单板的该方向上再次应用流策略；指定单板在某方向应用流策略后，也不能在全局的该方向上再次应用流策略。

在全局应用，系统对进入设备的所有匹配流分类规则的入方向或出方向报文流实施策略控制；在单板应用，系统对进入该单板的所有匹配流分类规则的入方向或出方向报文流实施策略控制。

缺省情况下，没有应用任何流策略，undo traffic-policy [policy-name] global { inbound | outbound } [slot slot-id] 命令用来删除在全局应用的流策略。

【示例 1】创建流策略p1并在该策略下关联已创建的流分类c1和流行为b1，然后在全局入方向上应用该策略。

<HUAWEI> system-view

[HUAWEI] traffic policy p1

[HUAWEI-trafficpolicy-p1] classifier c1 behavior b1

[HUAWEI-trafficpolicy-p1] quit

[HUAWEI] traffic-policy p1 global inbound

2．在接口上应用流策略

在接口上应用流策略是在具体接口视图下进行配置的。每个接口的每个方向上只能应用一个流策略，但同一个流策略可以同时应用在不同接口的不同方向。应用后，系统对流经该接口并匹配流分类中规则的入方向或出方向报文实施策略控制。但是流策略对VLAN 0的报文不生效。

建议不要在 Untagged 类型接口出方向上应用包含有 remark 8021p、remark cvlan-id、remark vlan-id等行为的流策略，否则，可能导致报文内容出错。不同S系列所使用的配置命令也有所不同，下面分别予以介绍。

（1）在 S2700-52P-EI/2700-52P-PWR-EI/2710SI/3700SI/3700EI 系列交换机上。使用traffic-policy policy-name inbound命令在接口的入方向应用流策略。缺省情况下，没有应用任何流策略，可用 undo traffic-policy inbound命令取消在接口入方向上应用流策略。

（2）在其他 S 系列交换机（ **S2700SI** 交换机不支持流策略应用 ）上。使用traffic-policypolicy-name { inbound | outbound }命令在接口的出方向或者入方向应用流策略。但只有在流行为中配置了流镜像功能，才可以将流策略应用于出方向，即出方向流策略只支持流镜像功能 。

缺省情况下，接口上没有应用任何流策略，可用 undo traffic-policy { inbound |outbound }命令取消在接口入方向或出方向上应用流策略。

【示例 2】创建流策略p1并在该策略下关联已创建的流分类c1和流行为b1，然后在GE0/0/1接口入方向上应用该策略。

<HUAWEI> system-view

[HUAWEI] traffic policy p1

[HUAWEI-trafficpolicy-p1] classifier c1behavior b1

[HUAWEI-trafficpolicy-p1] quit

[HUAWEI] interface gigabitethernet 0/0/1

[HUAWEI-GigabitEthernet0/0/1] traffic-policy p1 inbound

3．在VLAN上应用流策略

在VLAN上应用流策略必须在对应的VLAN视图下进行配置。应用后，系统对属于该VLAN并匹配流分类中规则的入方向报文实施策略控制。但是如果匹配到VLAN 0报文，则流策略不生效。同样不同S系列交换机所使用的配置命令有所不同，下面分别予以介绍。

（1）在 S2700/3700 系列交换机中（**S2700SI** 交换机不支持流策略应用 ）。在系统视图下使用traffic-policy policy-name global inbound命令在加入了对应VLAN的接口入方向上应用流策略。缺省情况下，没有应用任何流策略，可用undo traffic-policy inbound命令取消在加入了对应VLAN的接口入方向上应用流策略。

（2）在其他 S 系列交换机（ **S2700SI** 交换机不支持流策略应用 ）上。使用traffic-policypolicy-name { inbound | outbound }命令在加入了对应VLAN的接口入方向或者出方向应用流策略。缺省情况下，没有应用任何流策略，可用 undo traffic- policy { inbound | outbound }命令取消在加入了对应VLAN的接口入或出方向上应用流策略。

【示例 3】创建流策略p1并在该策略下关联已创建的流分类c1和流行为b1，然后在VLAN 100的入方向上应用该策略。

<HUAWEI>system-view

[HUAWEI] traffic policy p1

[HUAWEI-trafficpolicy-p1] classifier c1 behavior b1

[HUAWEI-trafficpolicy-p1] quit

[HUAWEI] vlan 100

[HUAWEI-vlan100] traffic-policy p1 inbound

11.4.5 基于复杂流分类的优先级重标记配置示例

本示例拓扑结构如图11-7所示，企业分支机构1和企业分支机构2通过Switch连接到外部网络设备，其中企业分支机构1属于VLAN100，企业分支机构2属于VLAN200。现希望分支机构1上送的数据报文能够得到更好的QoS保证，实现差分服务。

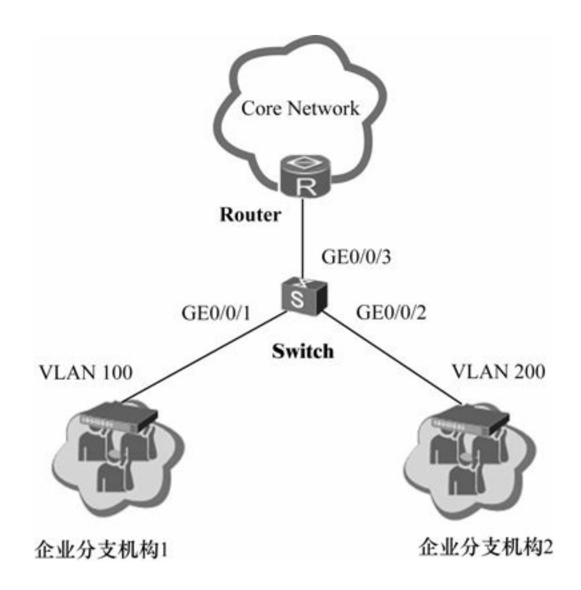


图11-7 基于复杂流分类的优先级重标记配置示例拓扑结构

1．基本配置思路

本示例要区分不同分支机构的 VLAN 报文优先级，所以可采用 QoS 流策略中的重标记报文的802.1p优先级的方式实现差分服务。根据QoS流策略的“定义流分类”、“定义流行为”、“创建流策略”和“应用流策略”4 个配置任务，可得出本示例的基本配置思路如下。

（1）在Switch上创建VLAN，并配置各接口类型为trunk，实现企业分支机构能通过Switch访问网络。

（2）在Switch上配置流分类，实现基于VLAN ID对报文进行分类。

（3）在 Switch 上配置流行为，将企业分支机构 1 和企业分支机构 2 上送的报文的802.1p优先级分别重标记为4和2，实现企业分支机构1的优先级高于企业分支机构2。

（4）在 Switch 上配置流策略，绑定已经配置好的流行为和流分类，并应用到接口GE0/0/1和GE0/0/2的入方向上，实现差分服务。

2．具体配置步骤

（1）在Switch上创建VLAN100和VLAN200。

<HUAWEI>system-view

[HUAWEI] sysname Switch

[Switch] vlan batch 100 200

（2）配置GE0/0/1、GE0/0/2和GE0/0/3接口的类型为Trunk，并将GE0/0/1接口加入VLAN100，将GE0/0/2接口加入VLAN200，GE0/0/3接口加入VLAN100和VLAN200。

[Switch] interfacegigabitethernet 0/0/1

[Switch-GigabitEthernet0/0/1] port link-type trunk

[Switch-GigabitEthernet0/0/1] port trunk allow-pass vlan 100

[Switch-GigabitEthernet0/0/1] quit

[Switch] interfacegigabitethernet 0/0/2

[Switch-GigabitEthernet0/0/2] port link-type trunk

[Switch-GigabitEthernet0/0/2] port trunk allow-pass vlan 200

[Switch-GigabitEthernet0/0/2] quit

[Switch] interface gigabitethernet 0/0/3

[Switch-GigabitEthernet0/0/3] port link-type trunk

[Switch-GigabitEthernet0/0/3] port trunk allow-pass vlan 100 200

[Switch-GigabitEthernet0/0/3] quit

（3）在Switch上定义并配置流分类c1、c2，对来自企业分支机构的报文按照其VLAN ID进行分类。

[Switch] traffic classifier c1 operator and

[Switch-classifier-c1] if-match vlan-id 100

[Switch-classifier-c1] quit

[Switch] traffic classifier c2 operator and

[Switch-classifier-c2] if-match vlan-id 200

[Switch-classifier-c2] quit

（4）在 Switch 上定义并配置流行为 b1、b2，分别重标记分支机构 1 和分支机构 2的VLAN报文的802.1p优先级为4和2。

[Switch] traffic behaviorb1

[Switch-behavior-b1] remark 8021p 4

[Switch-behavior-b1] quit

[Switch] traffic behaviorb2

[Switch-behavior-b2] remark 8021p 2

[Switch-behavior-b2] quit

（5）在 Switch 上创建流策略 p1，将前面定义的流分类和对应的流行为进行绑定，并将流策略应用到GE0/0/1和GE0/0/2接口的入方向上，对报文进行重标记。

[Switch] traffic policy p1

[Switch-trafficpolicy-p1] classifier c1 behaviorb1

[Switch-trafficpolicy-p1] classifier c2 behaviorb2

[Switch-trafficpolicy-p1] quit

[Switch] interface gigabitethernet 0/0/1

[Switch-GigabitEthernet0/0/1] traffic-policy p1 inbound

[Switch-GigabitEthernet0/0/1] quit

[Switch] interface gigabitethernet 0/0/2

[Switch-GigabitEthernet0/0/2] traffic-policy p1 inbound

```
[Switch-GigabitEthernet0/0/2] quit
```

配置好后，可以通过display traffic classifier user-defined命令查看流分类的配置信息，验证配置结果，具体如下所示。

```
<Switch>display traffic classifier user-defined
  User Defined Classifier Information:
  Classifier: c2
    Operator: AND
    Rule(s) : if-match vlan-id 200
  Classifier: c1
    Operator: AND
    Rule(s) : if-match vlan-id 100
Total classifier number is 2
```

也可以通过指定具体的流策略名查看流策略的配置信息，具体如下所示。

```
<Switch>display traffic policy user-defined p1
  User Defined Traffic Policy Information:
  Policy: p1
    Classifier: c1
      Operator: AND
        Behavior: b1
          Remark:
            Remark 8021p 4
    Classifier: c2
      Operator: AND
        Behavior: b2
          Remark:
            Remark 8021p 2
```

11.4.6 基于复杂流分类的流量统计配置示例

本示例拓扑结构如图 11-8 所示，PC1 的 MAC 地址为 0000-0000-0003，它连接在Switch的GE0/0/1接口上。现希望Switch对源MAC为0000-0000-0003的报文进行流量统计。

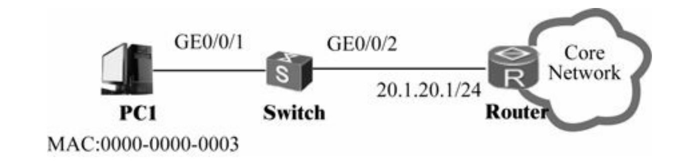


图11-8 流量统计配置示例拓扑结构

1．基本配置思路分析

本示例采用包含流量统计行为的 QoS 流策略方式实现流量统计（还是 QoS 流策略的四大配置任务进行的），具体配置思路如下。

（1）配置各接口，实现Switch与Router、PC1互通。

（2）配置二层ACL规则，匹配源MAC为0000-0000-0003的报文。

（3）定义流分类，实现基于上述ACL规则对报文进行分类。

（4）定义流行为，实现对满足规则的报文进行流量统计。

（5）创建流策略，绑定上述流分类和流行为，并应用到GE0/0/1接口的入方向，实现对该接口收到的源MAC为0000-0000-0003的报文进行流量统计。

2．具体配置步骤

（1）在Switch上创建VLAN20。

```
<HUAWEI>system-view
[HUAWEI] sysname Switch
[Switch] vlan 20
[Switch-vlan20] quit
```

（2）配置GE0/0/1接口为Access类型接口， GE0/0/2接口为Trunk类型接口，并将GE0/0/1和GE0/0/2加入VLAN20。

```
[Switch] interface gigabitethernet 0/0/1
[Switch-GigabitEthernet0/0/1] port link-type access
[Switch-GigabitEthernet0/0/1] port default vlan 20
[Switch-GigabitEthernet0/0/1] quit
[Switch] interfacegigabitethernet 0/0/2
[Switch-GigabitEthernet0/0/2] port link-type trunk
[Switch-GigabitEthernet0/0/2] port trunk allow-pass vlan 20
[Switch-GigabitEthernet0/0/2] quit
```

（3）创建VLANIF20，并配置IP地址20.1.20.2/24（需要配置Router与Switch对接的接口IP地址在同一网段，如20.1.20.1/24）。

```
[Switch] interface vlanif20
[Switch-Vlanif20] ip address 20.1.20.2 24
[Switch-Vlanif20] quit
```

（4）在Switch上创建编码为4000的二层ACL，匹配源MAC为0000-0000-0003的报文。

```
[Switch] acl 4000
[Switch-acl-L2-4000] rule permit source-mac0000-0000-0003 ffff-ffff-ffff
[Switch-acl-L2-4000] quit
```

（5）在Switch上定义流分类 c1，匹配规则为ACL 4000。

```
[Switch] traffic classifier c1 operator and
[Switch-classifier-c1] if-match acl 4000
[Switch-classifier-c1] quit
```

（6）在Switch上定义流行为b1，并配置流量统计行为。

```
[Switch] traffic behaviorb1
[Switch-behavior-b1] statistic enable
[Switch-behavior-b1] quit
```

（7）在Switch上创建流策略p1，将流分类和对应的流行为进行绑定。

```
[Switch] traffic policy p1
[Switch-trafficpolicy-p1] classifier c1 behaviorb1
[Switch-trafficpolicy-p1] quit
```

（8）将流策略p1应用到GE0/0/1接口入方向。

```
[Switch] interfacegigabitethernet 0/0/1
[Switch-GigabitEthernet0/0/1] traffic-policy p1 inbound
[Switch-GigabitEthernet0/0/1] quit
```

配置好后，可通过display traffic classifier user-defined命令查看流分类的配置信息，验证配置结果，通过display traffic policy user-defined命令查看流策略的配置信息，具体如下。

```
<Switch>display traffic classifier user-defined
  User Defined Classifier Information:
  Classifier: c1
    Operator: AND
    Rule(s) : if-match acl 4000
Total classifier number is 1
<Switch>display traffic policy user-defined p1
  User Defined Traffic Policy Information:
  Policy: p1
    Classifier: c1
      Operator: AND
        Behavior: b1
          statistic:enable
```

可通过display traffic policy statistics interface命令查看接口上的流量统计信息，具体如下。

```
<Switch>display traffic policy statistics interfacegigabitethernet0/0/1 inbound
Interface: GigabitEthernet0/0/1
Traffic policy inbound: p1
Rule number: 1
Current status: OK!
------------------------------------------------------------------
Board : 0
Item            Packets          Bytes
------------------------------------------------------------------
Matched              0             -
  +--Passed          0             -
  +--Dropped         0             -
    +--Filter        0             -
    +--CAR           0             -
```

11.4.7 基于复杂流分类的报文过滤配置示例

本示例拓扑结构如图11-9所示，企业用户通过SwitchA的GE0/0/2接口连接到外部网络设备。不同业务的

报文在 LSW 侧使用 802.1p 优先级进行标识，当报文从 GE0/0/2接口到达外部网络时，用户希望能够对数据业务报文进行过滤，优先保证话音和视频业务的业务体验。

1．基本配置思路分析

本示例是采用802.1p优先级过滤的流行为来实现报文过滤的，具体配置思路如下。

（1）配置各接口，实现企业用户能通过SwitchA访问外部网络。

（2）定义流分类，实现基于802.1p优先级对报文进行分类。

（3）定义流行为，实现对满足规则的报文进行禁止或允许行为。

（4）创建流策略，绑定上述定义的流分类和流行为，并应用到GE0/0/1接口的入方向，实现报文过滤。

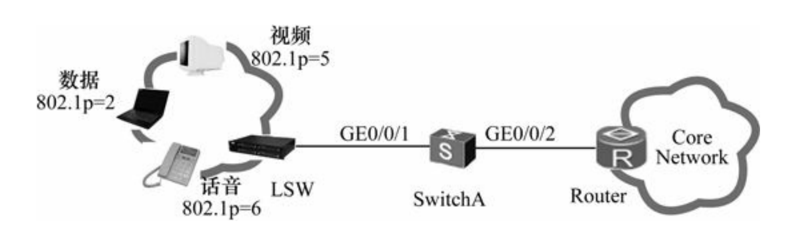


图11-9 报文过滤配置示例拓扑结构

2．具体配置步骤

（1）在Switch上创建VLAN10。

<HUAWEI>system-view

[HUAWEI] sysname SwitchA

[SwitchA] vlan 10

[SwitchA-vlan10] quit

（2）配置SwitchA上接口GE1/0/1和GE0/0/1为Trunk类型接口，并加入VLAN10（需要同时配置LSW与SwitchA对接的接口为Trunk类型，并加入VLAN10）。

[SwitchA] interface gigabitethernet 0/0/1

[SwitchA-GigabitEthernet0/0/1] port link-type trunk

[SwitchA-GigabitEthernet0/0/1] port trunk allow-pass vlan 10

[SwitchA-GigabitEthernet0/0/1] quit

[SwitchA] interface gigabitethernet 0/0/2

[SwitchA-GigabitEthernet0/0/2] port link-type trunk

[SwitchA-GigabitEthernet0/0/2] port trunk allow-pass vlan 10

[SwitchA-GigabitEthernet0/0/2] quit

（3）创建VLANIF10，并为VLANIF10配置IP地址192.168.2.1/24（需要同时配置Router与SwitchA对接的接口IP地址在同一网段，如192.168.2.2/24）。

[SwitchA] interface vlanif10

[SwitchA-Vlanif10] ip address 192.168.2.1 24

[SwitchA-Vlanif10] quit

（4）在SwitchA上定义并配置流分类c1、c2、c3，对报文按照802.1p优先级进行分类。

[SwitchA] traffic classifier c1

[SwitchA-classifier-c1] if-match 8021p 2

[SwitchA-classifier-c1] quit
[SwitchA] traffic classifier c2
[SwitchA-classifier-c2] if-match 8021p 5
[SwitchA-classifier-c2] quit
[SwitchA] traffic classifier c3
[SwitchA-classifier-c3] if-match 8021p 6
[SwitchA-classifier-c3] quit

（5）在SwitchA上定义流行为b1，并禁止行为，定义流行为b2和b3，并允许行为。

[SwitchA] traffic behavior b1
[SwitchA-behavior-b1] deny
[SwitchA-behavior-b1] quit
[SwitchA] traffic behaviorb2
[SwitchA-behavior-b2] permit
[SwitchA-behavior-b2] quit
[SwitchA] traffic behavior b3
[SwitchA-behavior-b3] permit
[SwitchA-behavior-b3] quit

（6）在SwitchA上创建流策略p1，将流分类和对应的流行为进行绑定并将流策略应用到GE0/0/1接口的入方向上，对报文进行过滤。

[SwitchA] traffic policy p1
[SwitchA-trafficpolicy-p1] classifier c1 behaviorb1
[SwitchA-trafficpolicy-p1] classifier c2 behaviorb2
[SwitchA-trafficpolicy-p1] classifier c3 behaviorb3
[SwitchA-trafficpolicy-p1] quit
[SwitchA] interfacegigabitethernet 0/0/1
[SwitchA-GigabitEthernet0/0/1] traffic-policy p1 inbound
[SwitchA-GigabitEthernet0/0/1] quit

配置好后，可通过display traffic classifier user-defined命令查看流分类的配置信息验证配置结果，具体如下。从输出信息可以看出每个分类中的各个流分类规则。

```
<SwitchA>display traffic classifier user-defined
  User Defined Classifier Information:
    Classifier: c2
      Operator: AND
      Rule(s) : if-match 8021p 5
    Classifier: c3
      Operator: AND
      Rule(s) : if-match 8021p 6
    Classifier: c1
      Operator: AND
      Rule(s) : if-match 8021p 2
```

Total classifier number is 3

可通过display traffic-policy applied-record命令查看流策略的应用信息，具体如下。从输出信息中可以看出流策略在GE0/0/1接口上应用是成功的。

```
<Switch>display traffic-policy applied-record p1
-------------------------------------------------
Policy Name:       p1
Policy Index:      3
     Classifier:c1    Behavior:b1
     Classifier:c2    Behavior:b2
     Classifier:c3    Behavior:b3
-------------------------------------------------
*interface GigabitEthernet0/0/1
   traffic-policy p1 inbound
     slot 0    :    success
-------------------------------------------------
Policy total applied times: 1.
```

第12章　IP组播基础及工作原理

IP组播在一些多用户定向发送的网络应用中使用非常普遍，如远程多媒体会议、远程教学、视频点播、定向电子商务，以及ISP的IPTV（网络电视）等。而这些应用又是目前最热门的互联网应用，在大多数公司中都有一些这类应用，所以学好IP组播基础知识及配置与管理方法，对于网络职业人士来说是非常必需的，特别是想成为专业、高薪的网络工程师们。

因为IP组播就像IP单播一样是一个相对独立的领域，所以涉及的知识面非常广，所包含的协议也非常多，如三层的IGMP、MLD、PIM、MSDP、MBGP和IGMP SMM Mapping等，二层的有IGMP Snooping、IGMP Snooping Proxy、MLD Snooping、MLD Snooping Proxy和IGMP SMM Snooping Mapping等。本章先单独介绍与IP组播相关的基础知识，组播路由和转发原理、各种IP组播协议的主要功能及工作原理，以及各种组播协议的主要应用。在下章将具体介绍华为S系列交换机中的各种IP组播应用配置与管理方法。

12.1 IP组播基础

随着Internet的不断发展，网络中交互的各种数据、话音和视频信息越来越多，同时新兴的电子商务、网上会议、网上拍卖、视频点播、远程教学等服务也在逐渐兴起。这些服务大多符合点对多点的模式，对信息安全性、有偿性、网络带宽提出了较高的要求。

作为IP传输3种方式之一，IP组播通信指的是IP报文从一个源发出，而被转发到一组特定的接收者。相较于传统的单播和广播，I P组播可以有效地节约网络带宽、降低网络负载，所以在IPTV、实时数据传送和多媒体会议等诸多方面都有广泛的应用。

12.1.1 IP网络的3种数据传输方式

IPv4协议定义了三种IP 数据包的传输方式：单播（unicast）、广播（broadcast）和组播（multicast）。下面首先对这三种包传输方式进行比较式地介绍，从中可以看出组播方式的优越性。

1．单播方式的数据传输过程

单播用于发送数据包到单个目的地，且每发送一份单播报文都使用一个单播IP地址作为目的地址。这是最常见的IP传输方式，是一种点对点传输方式。采用单播方式时，系统为每个需求该数据的用户单独建立一条数据传送通路，并为该用户发送一份独立的副本数据。

如图 12-1 所示，假设用户 C（HostC）需要从数据源（Source）获取数据，则数据源必须和用户C的设备建立单独的传输通道。由于网络中传输的数据量和要求接收该数据的用户量成正比，因此当需要相同数据的用户数量很庞大时，数据源主机就必须要将多份内容相同的数据发送给用户。这样一来，网络带宽将成为数据传输中的瓶颈，所以不利于数据规模化发送。

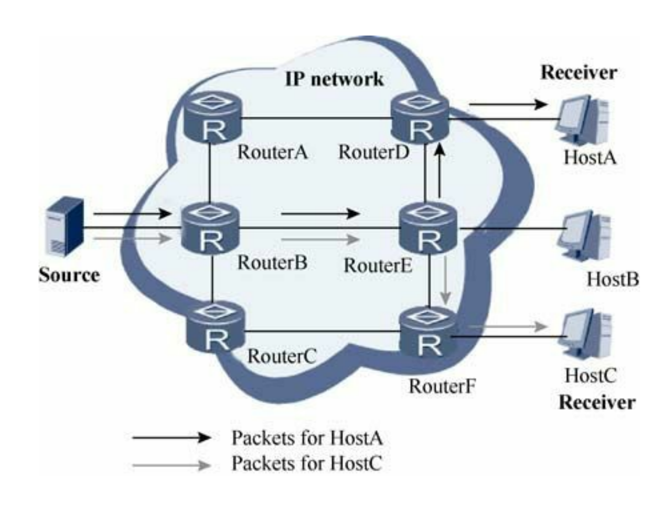


图12-1 单播方式传输数据示意图

2．广播方式的数据传输过程

广播是指发送数据包到同一广播域或子网内的所有设备的一种数据传输方式，是一 种点对多点传输方式。如果采用广播方式，系统会为网络中所有用户传送一个数据副本，不管他们是否需要，任何用户都会接收到广播来的数据。

如图12-2所示假设用户A、C需要从数据源获取数据，则数据源通过路由器广播该数据，但这时网络中本来不需要接收该数据的用户B 也同样接收到该数据，这样不仅信息的安全性得不到保障，而且会造成同一网段中信息泛滥。由此可见，该传输方式不利于与特定对象进行数据交互，并且浪费了大量的带宽。

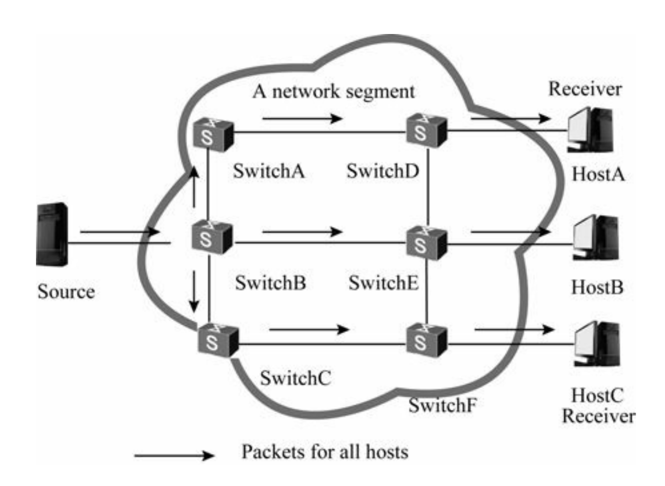


图12-2 广播方式传输数据示意图

3．组播方式传输数据

通过前面的介绍可以看出，单播方式适合用户较少的网络，而广播方式适合用户需求普遍相同的网络。但当网络中需求某数据的用户量不确定时，单播和广播方式效率很低，而且广播方式安全性差，无法控制数据的发送。

IP组播技术的出现及时解决了以上这些问题，也是一种点对多点传输方式。当网络中的某些用户需要特定数据时，组播数据发送者（即组播源）仅发送一次数据，借助组播路由协议为组播数据包建立组播分发树，被传递的数据到达距离用户端尽可能近的节点后才开始复制和分发。

如图12-3所示，假设用户A、C需要从数据源获取数据，为了将数据顺利地传输给真正需要该数据的用户，需要将用户A、C组成一个接收者集合（就是组播组），由网络中各路由器根据该集合中各接收者的分布情况进行数据转发和复制，最后准确地传输给实际需要的接收者A和C。

综上所述，相比单播传输方式，组播传输方式由于被传递的信息在距信息源尽可能远的网络节点才开始被复制和分发，所以用户的增加不会导致信息源负载的加重以及网络资源消耗的显著增加。相比广播传输方式，组播传输方式由于被传递的信息只会发送给需要该信息的接收者，所以不会造成网络资源的浪费，并能提高信息传输的安全性。

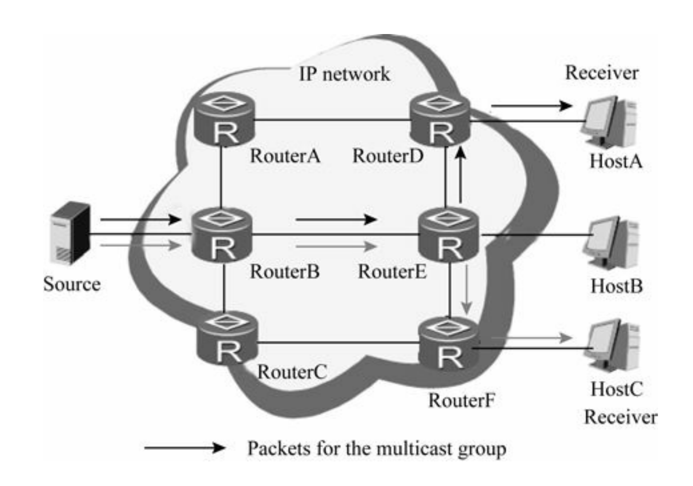


图12-3 组播方式传输数据示意图

12.1.2 组播基本概念

组播传输的特点是一点发出，多点接收。图12-3所示为组播的传输模型示意图，网络中存在信息发送源Source，感兴趣的用户HostA和HostC提出信息需求，Source发出的数据只有HostA和HostC会接收到。

在组播通信中，需要理解以下几个重要的基本概念。

（1）组播组：用组播IP地址进行标识的一个集合，是一个组播成员的集合，各组播成员共享这一个组播组IP地址。就相当于在本书前面介绍的iStack堆叠和CS集群中，各成员交换机共享使用同一个管理IP地址一样。但要注意，组播成员自己在**IP**协议中配置的**IP**地址不是组播**IP**地址，仍是单播**IP**地址 。任何用户主机（或其他接收设备），加入一个组播组就成为了该组成员，可以识别并接收发往该组播组的组播数据。

（2）组播源：以组播组IP地址为目的地址（组播源配置的也是单播**IP**地址 ），发送IP报文的信源称为组播源。但组播源通常不需要加入组播组，否则自己接收自己发送出去的数据了。图 12-3 中的 Source 就是一个组播源。一个组播源可以同时向多个组播组发送数据，多个组播源也可以同时向一个组播组发送报文。

（3）组播组成员：所有加入某组播组的主机便成为该组播组的成员，如图12-3中的HostA 和 HostC。组播组中的成员是动态的，主机可以在任何时刻加入或离开组播组。组播组成员可以广泛地分布在网络中的任何地方。

（4）组播路由器：支持三层组播功能的路由器或三层交换机（它们不是组播组成员），如图12-3中的各个Router。组播路由器不仅能够提供组播路由功能，也能够在与用户连接的末梢网段上提供组播组成员的管理功能。

12.1.3 典型IP组播模型

根据对组播源处理方式的不同，IP组播模型有下列3种：ASM（Any-Source Multicast，任意源组播）、SFM（Source-Filtered Multicast，过滤源组播）和 SSM（Source-Specific Multicast，指定源组播）。

1．ASM模型

简单地说，ASM（任意源组播）模型就是任意源都可以成为组播源。由此可知，ASM模型中的组播源是不限定的，任意一个发送者都可以成为组播源，然后向某组播组地址发送数据，显然安全性较差。接收者通过加入对应的组播组就可以获得发往该组播组的任意组播数据，而且接收者无法预先知道组播源的位置，但可以在任意时间加入或离开该组播组。

为了提高安全性，可以在路由器上配置针对组播源的过滤策略，允许或禁止来自某些组播源的报文通过。最终从接收者角度看，数据是经过筛选的。

ASM 模型中组播组可以使用的组播 IP 地址为 224.0.1.0～231.255.255.255、233.0.0.0～238.255.255.255。但要求组播组地址必须整个组播网络中唯一。“唯一”指的是同一时刻一个 ASM 组播组地址只能被一种组播应用使用。如果有两种不同的应用程序使用了同一个 ASM 组播组地址发送数据，它们的接收者会同时收到来自两个源的数据。这样一方面会导致网络流量拥塞，另一方面也会给接收者主机造成困扰。

2．SFM模型

SFM（过滤源组播）模型继承了 ASM 模型，从发送者角度来看两者的组播组成员关系完全相同，也可以是任意组播源 。但是，在SFM模型中组播上层应用软件可以根据收到的组播包的源IP地址进行过滤，允许或禁止来自某些组播源的包通过。这样一来，接收者就可以只接收允许通过的组播源发来的组播数据。即SFM在ASM的基础上添加了组播源过滤策略。

3．SSM模型

在现实生活中，用户可能只对某些组播源发送的组播数据感兴趣，而不愿接收其他源发送的数据。SSM（指定源组播）模型就是一种为用户提供能够在客户端指定组播源的传输服务。

SSM模型与ASM模型的根本区别在于：SSM模型中的接收者已经通过其他手段预先知道了所需接收组播数据的组播源的具体位置，限定了可接收的组播源。然后，SSM模型使用与 ASM/SFM 模型不同的组播组地址范围（为 232.0.0.0～232.255.255.255）直接在接收者和其指定的组播源之间建立专用的组播转发树。

12.1.4 IP组播地址

由于组播数据的接收者是一个组播组内的多个主机，因此，需要面对数据源该将数据发往何处、目的地址如何选取的问题。这些问题简而言之就是组播寻址。与单播中的IP寻址或者MAC寻址一样，为了让组播源和组播组成员进行通信，需要提供网络层组播地址，即IP组播地址。同时必须存在一种技术将IP组播地址映射为链路层MAC组播地址。下面分别介绍这两种组播地址。

【经验之谈 】 千万不要认为在IP组播中所有组播设备上的IP地址都是使用组播地址。实际上只有组播组IP地址是组播IP地址，而像组播源、接收主机的IP地址仍是单播IP地址。

1．三层组播IP地址

根据 IANA（Internet Assigned Numbers Authority，因特网编号授权委员会）规定， IP地址分为五类，即A类、B类、C类、D类和E类。单播包按照网络规模大小分别使用A、B、C三类IP地址。组播包的目的地址使用D类IP地址，D类地址不能出现在IP包的源IP地址字段（也就是不能作为组播源地址，换言之，组播源的IP地址仍是单播地址）。E类地址保留在今后使用。

在单播数据传输过程中，一个数据包传输的路径是从源地址路由到目的地址，利用“逐跳”（hop-by-hop）的原理在IP网络中传输。然而在IP组播环境中，数据包的目的地不是一个，而是一组，形成组地址

（可以理解为所有接收者的单播地址与一个组播组地址形成了映射关系）。所有的数据接收者都加入一个组内，并且一旦加入之后，流向该组地址的数据立即向接收者传输，组中的所有成员都能接收到数据包，这个组就是“组播组”。

组播组具有以下几个特点。

（1）组播组中的成员是动态的，主机可以在任何时刻加入和离开组播组。

（2）组播组可以是永久的也可以是临时的。

（3）由IANA分配组播地址的组播组称为永久组播组（又称保留组播组）。

对于永久组播组，要注意以下几点。

（1）永久组播组的IP地址保持不变，但组中的成员构成可以发生变化。

（2）永久组播组中成员的数量可以是任意的，甚至可以为零。

（3）那些没有保留下来供永久组播组使用的IP组播地址，可以被临时组播组使用。

D 类组播地址范围是 224.0.0.0～239.255.255.255，其中包括了很多地址，但不同地址段有不同用途，具体如表12-1所示。记住这个表中各个组播段的使用范围相当重要，这样就不会在配置组播网络中错误地使用了不该在特定环境下使用的组播地址。

表12-1 D类地址的范围及用途

D 类地址范围	用途
224.0.0.0～224.0.0.255	预留的组播组地址（也就是永久组地址），地址 224.0.0.0 保留不做分配，其他地址供路由协议使用
224.0.1.0～231.255.255.255 233.0.0.0～238.255.255.255	可用的 ASM（Any-Source Multicast，任意源组播模型）组播组地址，全网范围内有效
232.0.0.0～232.255.255.255	可用的 SSM（Source-Specific Multicast，指定源组播模型）组播组地址
239.0.0.0～239.255.255.255	本地管理组播地址，仅在特定的本地范围内有效

根据IANA的约定，224.0.0.0～224.0.0.255网段地址被预留给本地网络中的路由协议使用，常用的预留IP组播地址及用途说明如表12-2所示。

表12-2 预留IP组播地址及用途说明

D 类地址范围	用途
224.0.0.1	网段内所有主机和路由器（等效于广播地址）
224.0.0.2	所有组播路由器的地址
224.0.0.3	不分配
224.0.0.4	所有 DVMRP（Distance Vector Multicast Routing，距离矢量组播路由协议）路由器
224.0.0.5	所有 OSPF（Open Shortest Path First，最短路径优先）路由器
224.0.0.6	OSPF DR（Open Shortest Path First Ddesignated Router，最短路径优先指定路由器）
224.0.0.7	ST（Shared Tree，共享树）路由器
224.0.0.8	ST（Shared Tree，共享树）主机
224.0.0.9	RIPv2 路由器
224.0.0.11	活动代理
224.0.0.12	DHCP 服务器/中继代理
224.0.0.13	所有 PIM（Protocol Independent Multicast，协议无关组播）路由器
224.0.0.14	RSVP（Resource Reservation Protocol，资源预留协议）封装
224.0.0.15	所有 CBT（Core-Based Tree，有核树）路由器
224.0.0.16	指定 SBM（Subnetwork Bandwidth Management，子网带宽管理）
224.0.0.17	所有 SBMS
224.0.0.18	VRRP（Virtual Router Redundancy Protocol，虚拟路由冗余协议）
224.0.0.22	所有 IGMPv3 路由器

说明

与IANA为IP单播预留私有地址网段10.0.0.0/8等类似，IANA也为IP组播预留 了私有地址网段239.0.0.0/8（也就是这个地址段中的组播地址可以在局域网内使用）。这些地址属于管理范围地址，可以灵活地定义组播域范围，实现不同组播域之间的地址隔离，有助于相同组播地址在不同组播域内的重复使用而不会冲突。

2．二层以太网组播MAC地址

以太网传输单播IP包的时候，目的MAC地址使用的是接收者的MAC地址。但是在传输组播包时，传输目标不再是一个具体的接收者，而是一个成员不确定的组，所以对应也就需要使用组播MAC地址作为目的地址。

IANA规定，组播MAC地址的高 25位固定为 0000 0001 0000 0000 0101 1110 0，形成MAC地址25位前缀，MAC地址的低23位为组播IPv4地址的低23位。它们之间的映射关系如图12-4所示（组播IPv4地址中的低23位映射到组播MAC地址的低23位）。

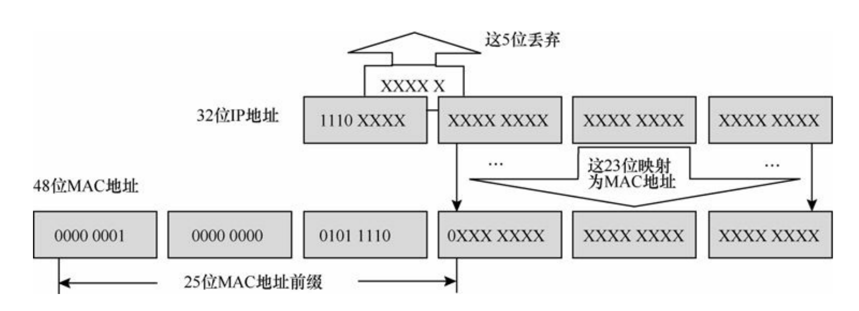


图12-4 IPv4组播地址到MAC组播地址的映射

由于IPv4组播地址的高4位是1110，代表组播标识，而低28位中只有23位被映射到MAC地址，这样IP地址中就会有5位数据丢失，直接的结果是出现了32（2$lt@sup$gt@5$lt@/sup$gt@）个IP组播地址映射到同一组播MAC地址上。

IPv6组播MAC地址的高16位为0x3333，低32位是从IPv6组播地址的低32位映射过来的。如图12-5所示的是IPv6组播地址FF1E::F30E:101的MAC地址映射举例。

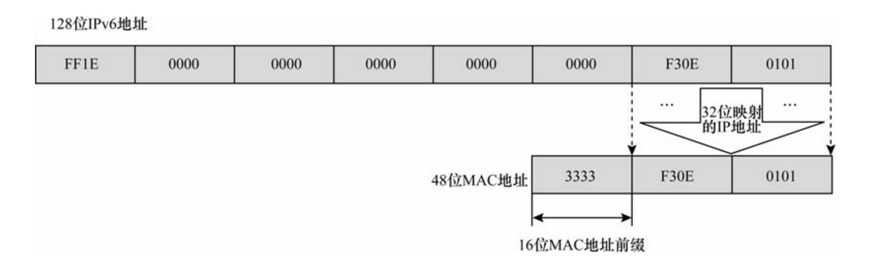


图12-5 IPv6组播地址的MAC地址映射示例

12.1.5 IP组播协议

要实现一套完整的组播服务，需要在网络各个位置部署多种组播协议相互配合，共同运作。但不同结构的组播网络所需使用的组播协议不完全一样。

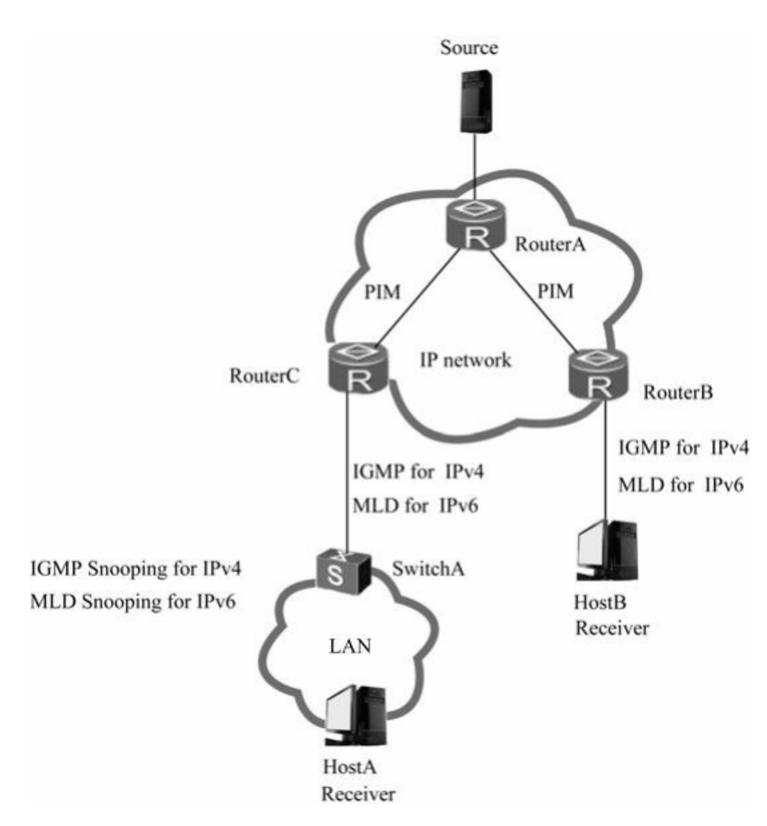
图12-6 是一个典型的单PIM 域组播网络示意图，整个组播网络是由路由器或三层交换机+二层交换机组成的，从中可以看出在这些组播设备上所运行的组播协议包括PIM （协议无关组播，同时有IPv4和IPv6版本）、IPv4网络的IGMP（因特网组管理协议）， IPv6网络中的MLD（组播监听器发现）、IPv4网络的

IGMP Snooping（因特网组管理协议嗅探），IPv6网络中的MLD Snooping（组播监听器发现嗅探）。

图12-7是一个跨PIM-SM域的组播网络示意图，它与图12-6所示的单PIM域组播网络相比，在运行的组播协议上仅需在PIM域边界组播路由器上多了一个实现跨PIM域连接的MSDP（组播源发现协议）。而图12-8是一个跨AS域组播网络示意图，它与图12-7所示的跨PIM域组播网络相比，在运行的组播协议又仅需在AS边界组播路由器上多了一个用于不同AS组播连接的MBGP（组播边界管理协议）。因为MBGP将在配套图书《华为路由器学习指南》中有详细介绍，故在此不再赘述。下面仅从各自的基本用途方面介绍 IGMP、IGMP Snooping、MLD、MLD Snooping、PIM和MSDP这六种组播协议。在本章后还将对它们的工作原理进行详细剖析。

说明

本章后面仅介绍IPv4组播网络的相关协议及配置与管理方法。



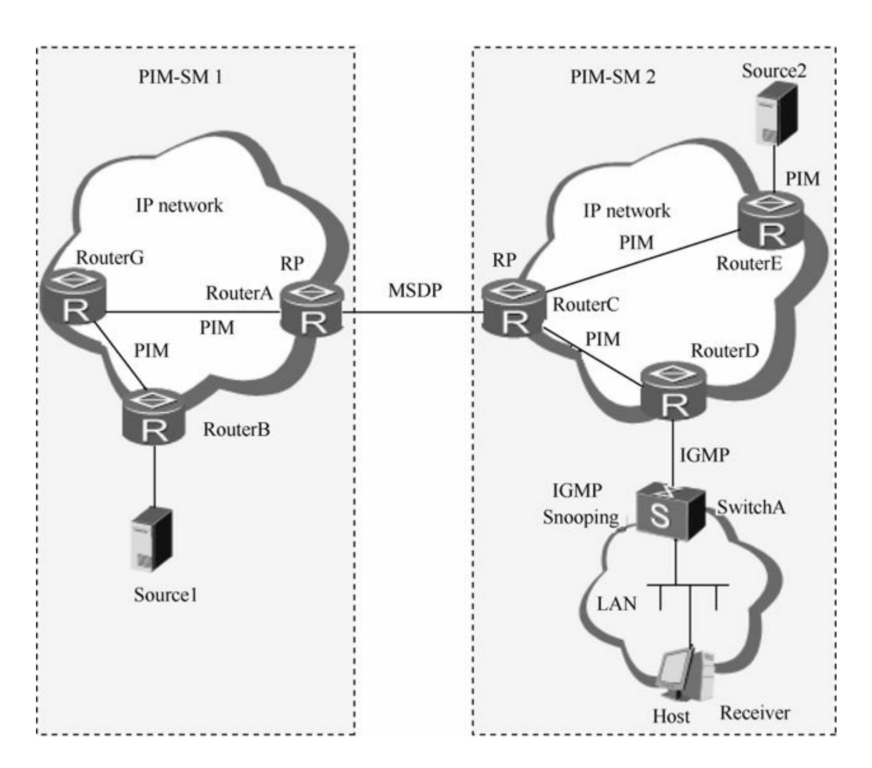


图12-7 跨PIM-SM域的组播网络示意图

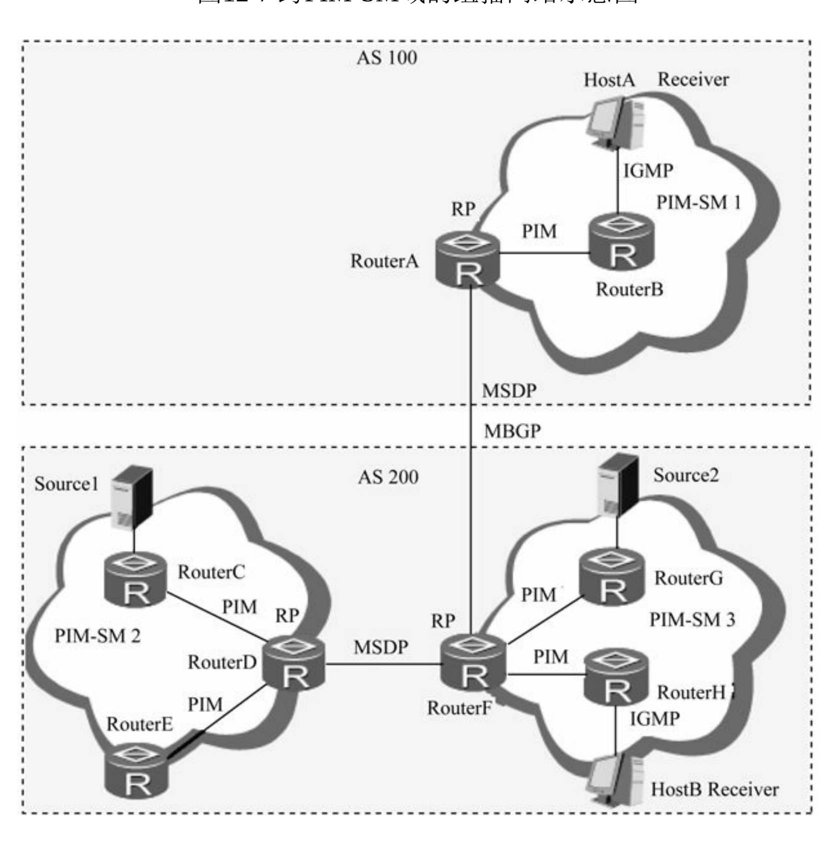


图12-8 跨AS域组播网络示意图

1．IGMP和MLD

在IP组播传输模型中，发送者不关心接收者所处的位置，只要将数据发送到约定的目的地址，剩下的工作就交给网络去完成。网络中的路由器设备必须收集接收者的信息，并按照正确的路径实现组播报文的转发和复制。

接收者信息的收集和管理的工作通过 IGMP（Internet Group Management Protocol，因特网组管理协议）或MLD（Multicast Listener Discovery，组播监听器发现）协议来完成的。其中，IGMP用于IPv4网络，MLD用于IPv6网络。用于为主机侧提供组播组成员动态加入与离开服务，为路由器侧提供组成员关系的维护与管理服务，同时与上层组播路由协议进行信息交互。

IGMP包含3个版本，分别是IGMPv1、IGMPv2和IGMPv3。新版本完全兼容旧版本。目前应用最广泛的是IGMPv2。在组播模型方面，3个版本都支持ASM模型；IGMPv3可以直接支持SSM模型，而IGMPv1和IGMPv2需要结合SSM-Mapping技术才能支持SSM模型。

在 IPv6 组播中使用 MLD 协议来替代 IGMP 协议，也是一种三层组播协议。MLD包含两个版本，分别是MLDv1和MLDv2。MLDv1的功能与IGMPv2相似；MLDv2的功能与IGMPv3相似。两个MDL版本都支持ASM模型；MLDv2可以直接支持SSM模型，而MLDv1也需要结合SSM-Mapping技术才能支持SSM模型。

2．IGMP Snooping和MLD Snooping

IGMP Snooping和MLD Snooping协议是运行在组播路由器和用户主机之间的二层交换机上的二层组播协议，配置在VLAN内。其中，IGMP Snooping用于 IPv4网络，MLD Snooping用于 IPv6网络，用来侦听路由器和主机之间发送的 IGMP、MLD报文建立组播数据的二层转发表，从而管理和控制组播数据在二层网络中的转发。

3．PIM和MSDP

组播报文转发路径的建立，有多种组播路由协议可以完成。目前应用广泛的是PIM （Protocol Independent Multicast，协议无关组播）协议。PIM是一种域内组播路由协议，当跨PIM域传递组播源信息时，需要MSDP（Multicast Source Discovery Protocol，组播源发现协议）支持；当跨AS域建立组播路由时则同时需要MSDP和MBGP（MultiProtocol Border Gateway Protocol，组播边界网关协议）支持。

PIM是用于 IPv4或 IPv6组播网络（对应 IPv4 PIM版本和 IPv6 PIM版本）中域内组播路由器之间的组播路由与转发，用来在自治系统 AS 内发现组播源并构建组播分发树，将信息传递到接收者。在一个小型网络中，所有的组播路由器都在一个PIM组播域内。它可以动态响应网络拓扑变化，维护组播路由表，并按照路由表项执行转发。PIM有两套独立的模式。

（1）DM（Dense Mode）：适用于小规模、接收者分布较为密集的情况，支持 ASM模型。

（2）SM（Sparse Mode）：适用于大规模、接收者分布较为稀疏的情况，同时支持ASM模型和SSM模型。

如图12-6所示，为了便于控制和管理组播资源（组播组、组播源和组播组成员）， 需要将组播资源在域间进行隔离，从而形成一个个隔离的PIM-SM域。图12-7所示为跨PIM-SM域的组播网络。由于PIM协议依赖于单播路由表 ，所以组播转发路径与单播转发路径是一致的。当组播源与接收者分布在不同的 AS 中时，需要跨 AS 建立组播转发树，如图12-8所示。此时可以部署MBGP协议，生成一张独立于单播路由的组播路由表，使组播数据通过组播路由表 进行传输。

MSDP 目前仅用于 IPv4 组播网络中域间组播路由器之间的域间组播源信息共享，但只对 ASM 服务模型有意义。它可以实现源所在域内的路由器将本地源信息传播给其他域内的路由器，以及不同域的路由器之间传递源信息。为了使不同的PIM-SM域之间组播数据能够互通，需要在域间部署MSDP协议。MSDP通过在各个 PIM-SM域之间建立MSDP对等体关系，对等体之间交互 SA（Source Active，源激活）消息来传递组播信息，从而实现接收者主机可以接收其他PIM-SM域的组播源数据。

12.2 IGMP的3个版本及各自工作原理

IGMP（因特网组管理协议）是 TCP/IP协议族中负责 IP组播成员管理的一个子协议，用来在IP主机和与其直接相邻的组播路由器之间（不是应用于多个组播路由器之间）建立、维护组播组成员关系。IGMP消息封装在IP报文中，其IP的协议号为2。

到目前为止，IGMP有 3个版本：IGMPv1（由RFC 1112定义）、IGMPv2（由RFC 2236定义）和IGMPv3（由RFC 3376定义）。IGMPv1中定义了基本的组成员查询和报告过程，IGMPv2在此基础上添加了查询器选举和组成员离开的机制，IGMPv3中增加的主要功能是成员可以指定接收或指定不接收某些组播源的报文。

IGMP的3个版本在演进过程中对协议报文的处理是向前兼容的，因此，尽管各个版本的协议报文格式不同，但是运行IGMP高版本的路由器可以识别低版本的成员报告。所有这 3 个版本的 IGMP 都支持 ASM 模型，IGMPv3 可以直接应用于 SSM 模型。而IGMPv1和IGMPv2则需要SSM-Mapping技术的支持。

12.2.1 IGMPv1工作原理

IGMPv1 是最初的版本，主要基于查询和响应机制来完成对组播组成员的管理。IGMPv1 主机可以通过发送加入消息加入直接相连的组播路由器上特定的组播组，但离开时不会发送离开信息（leave messages）。IGMPv1组播路由器使用基于超时的机制去发现其成员不关注的组。

IGMPv1报文有两种类型。

（1）普遍组查询报文（General Query）：是查询器主动 向共享网络上所有主机和路由器发送的查询报文，用于了解哪些组播组存在成员。

（2）成员报告报文（Report）：是主机为了响应普遍查询报文而被动 向组播路由器发送的，或者是主机主动向组播路由器发送的报告消息，用于应答普遍查询报文加入某个 组播组，或者主动申请加入某个组播组。

IGMPv1协议主要基于查询/响应机制完成组播组管理。当一个网段内有多个组播路由器时，只需要其中一台发送查询报文就足够了，此时需要选举出一个IGMP查询器。在 IGMPv1中，由组播路由协议 PIM选举出唯一的组播信息转发者（Assert Winner或DR）作为IGMPv1的查询器，负责该网段的组成员关系查询。

1．普遍组查询和响应机制

如图12-9所示（左图显示的是普遍查询报文的传递过程，右图显示的是成员报告报文的传递过程），组播网络中RouterA和RouterB连接主机网段，RouterA为查询器（由PIM路由器选举决定，具体将在本章后面介绍PIM协议时介绍）。在主机网段上有HostA、HostB、HostC三个接收者。现假设HostA和HostB想要接收发往组播组G1的数据，HostC想要接收发往组播组G2的数据（均需要事先要经过相应配置 ）。普遍组查询和响应过程如下。

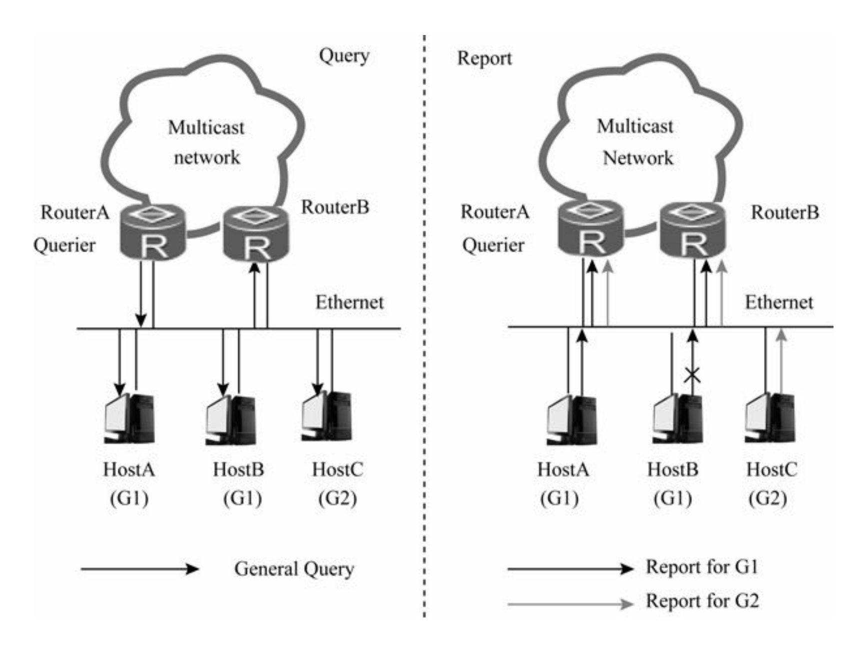


图12-9 普遍组查询和响应机制示意图

（1）IGMP查询器（RouterA）以目的地址224.0.0.1（这是一个永久组播地址，代表了同一网段内所有主机和路由器，相当于广播地址）向该网段中所有主机和路由器发送普遍组查询报文，以了解该网段中有哪些组播成员存在。普遍组查询报文是周期性发送的，发送周期可以通过命令配置，缺省情况下每隔60s发送一次。

（2）网段内所有主机和路由器都会接收到该查询报文，但只有已配置了对应组播组的成员才做出响应。为了避免多个成员主机同时发送响应的报告报文，在IGMP中规定，每个组播组会要求组播成员在发送报告报文前要启动一个随机定时器（为 0～10s），只有定时器超时后才会发送报告报文。因为HostA和HostB是组播组G1成员，所以均会在本地启动定时器 Timer-G1；HostC 是组播组 G2 的成员，同样会在本地启动定时器Timer-G2。

（3）当同一个组播组中的第一个成员启动的定时器超时后，会以224.0.0.1为目的地址发送针对该组播组的报告报文，在该网段中所有主机和组播路由器（包括IGMP查询 器）上都可以收到该报文，但只有查询器才会做出相应的响应。现假设 HostA 上的Timer-G1首先超时，它会向该网段发送目的地址也为224.0.0.1的报告报文，此时想加入组G1的HostB也可以侦听到HostA发送的报告报文，则停止自己的定时器Timer-G1，不再发送针对G1的报告报文，这样做的目的是可以减少网段上的流量。

同样，HostC 上的 Timer-G2 超时后也会向该网段发送报告报文，目的地址也为224.0.0.1。

（4）路由器接收到报告报文后，了解到本网段内存在组播组G1和G2的成员（但不需要了解具体所有成员，只需要知道哪些组播组中有成员 ），然后由路由器上运行的PIM协议生成对应的两个组播转发表项——（*，G1）和（*，G2），“*”代表任意组播源。这样，网络中一旦有组播组G1和G2的数据到达组播路由器，则将向该网段的对应组播 组中的所有成员主机转发。

2．新组成员加入

注意，这里所说的是有成员要加入新的组播组，并不是现有组播组中有新成员加入 。如图12-10所示，假设在网段上新接入一个主机HostD，想加入组播组G3（事先要通过配置），这是一个在PIM路由器还没建立的新的组播组。此时，该成员主机不会等待查询器的普遍组查询报文的到来，立即主动发送针对 G3 的报告报文（目的地址也为224.0.0.1）。查询器在收到报告报文后，了解到本网段内出现了新的组播组G3

的成员，会由上层PIM协议生成新的组播转发表项——（*，G3）。这样，当网络中一旦有G3的数据到达该PIM路由器后将向该网段的对应组播组中的所有成员主机转发。

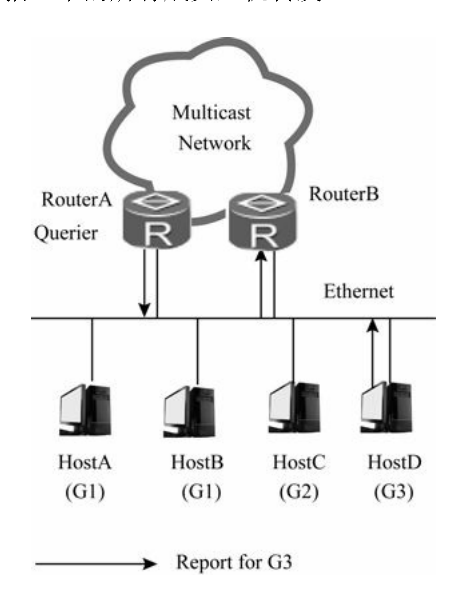


图12-10 新成员加入示意图

3．组成员离开

IGMPv1 没有专门定义离开组的报文。当主机离开组播组时不会再对后续的普遍组查询报文做出回应，纯粹通过普遍查询报文的最大响应定时器来确定某组播组中是否还有组播成员。假设图12-10中的HostC退出组播组G2，当收到普遍组查询报文时，HostC不再发送针对G2的报告报文进行响应。此时，由于网段上不存在组G2的其他成员，查询器永远不会收到G2的报告报文，会在一定时间（130s）后删除G2所对应的组播转发项。但如果是HostA退出组播组G1，则路由器不会感知到他的离开，因为G1中还有成员HostB，它会进行响应的。

12.2.2 IGMPv2的改进

与 IGMPv1 相比，IGMPv2 增加了独立的查询器选举机制（IGMPv1 中的查询器是组播路由协议选举指定路由器（DR）担当查询器的）和离开组机制，包含了离开信息，允许迅速向组播路由协议（如PIM）报告组成员终止情况，这对高带宽组播组或易变型组播组成员而言是非常重要的。

下面具体介绍IGMPv2新增的查询器选举机制和成员离开组机制。

1．查询器选举

IGMPv2使用独立的查询器选举机制，当共享网段上存在多个组播路由器时，IP地址最小的路由器成为查询器，当然这是由各个运行IGMP协议的路由器之间自动选举的，但可以通过IP地址设置间接指定查询器。下面以图12-11中的两台IGMP路由器为例介绍IGMP查询器的选举机制（左图为查询器选举初始状态下的普遍查询报文传递情况，右图为查询选举后的普遍查询报文传递情况）。

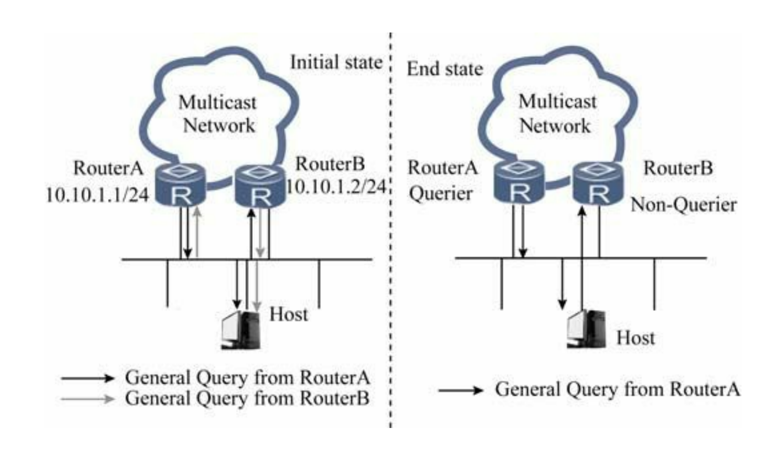


图12-11 查询器选举示意图

（1）所有 IGMPv2 路由器在初始时都认为自己是查询器，向本地网段（本示例为10.10.1.0/24）内的所有主机和路由器发送普遍组查询报文。

（2）其他路由器在收到某路由器发来的普遍查询报文后，将报文的源IP 地址与自己的接口地址作比较。通过比较，IP 地址最小的路由器将成为查询器，其他路由器成为非查询器（Non-Querier）。本示例中，RouterA的接口地址小于RouterB，所以RouterA最终当选为查询器。

所有非查询器上都会启动一个定时器，即“其他查询器存在时间定时器”（Other Querier Present Timer）。在该定时器超时前，如果收到了来自查询器的查询报文，则重置该定时器；否则，就认为原查询器失效，并发起新的查询器选举过程。这就相当于像RIP、OSPF这些动态路由协议中用来决定邻居路由器是否有效的Hello定时器。

2．离开组机制

如图12-12所示，如果两路由器都运行了IGMPv2协议，现主机HostC想离开组播组G2，将会发生以下流程。

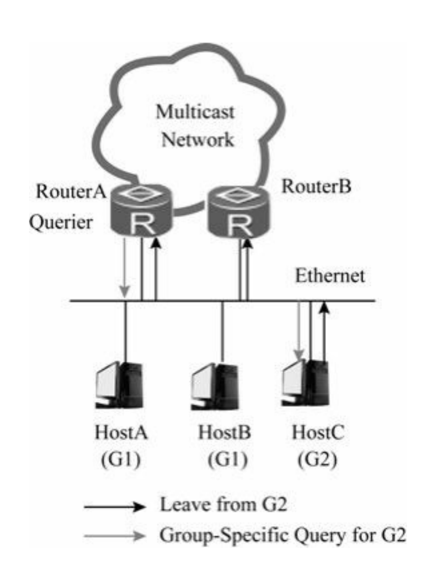


图12-12 离开组示意图

（1）在 HostC 离开组时向本地网段内所有组播路由器（目的地址为 **224.0.0.2**，这是一个代表本网段内所有组播路由器的永久组播地址 ）发送一个离开组播组G2的报文。

（2）当查询器收到离开报文后，会发送针对组播组G2的特定组查询报文（注意这里不是“普遍查询报文”），以查询本网段内是否还有其他组播组 G2 的成员。发送间隔和发送次数可以配置。缺省每隔1s发送一次，一共发送两次。同时启动组成员关系定时器 Timer-Membership=发送间隔x发送次数。

（3）如果本网段内不存在其他组播组G2的成员，则路由器不会收到针对组播组G2的报告报文，这时会在Timer-Membership定时器超时后，删除（*，G2）表项。这样，组播组G2的数据再到达路由器时，将不会再向该网段转发。

但在Timer-Membership定时器超时前有其他G2组播组成员以报告报文进行响应特定组查询报文，则表明该网段内还有组播组 G2 的其他成员，路由器继续维护该组成员关系。

12.2.3 IGMPv3的改进

IGMPv3 在兼容和继承 IGMPv1 和 IGMPv2 的基础上进一步增强了主机的控制能力，支持指定组播源/组播组功能，即主机在加入某组播组 G 的同时能够明确地要求接收或不接收某特定组播源 S 发出的组播信息。这主要是为了配合 SSM 模型发展起来的，提供了在报文中携带组播源信息的能力，使组播成员能加入指定源的组播组。

1．IGMPv3报文

IGMPv3报文也包含两大类：查询报文和报告报文。相较IGMPv2，其变化如下。

（1）查询报文中除了普遍组查询和特定组查询，新增了特定源组查询报文（Group-and-Source-Specific Query）。该报文由查询器向共享网段内特定组播组成员发送，用于查询该组成员是否愿意接收特定组播源发送的数据。特定源组查询通过在报文中携带一个或多个组播源地址来达到这一目的。

（2）报告报文不仅通知路由器主机要加入某组播组，并且可以指定只接收哪些组播源发往该组的数据。IGMPv3增加了针对组播源的过滤模式（INCLUDE/EXCLUDE），将组播组与源列表之间的对应关系简单地表示为（G，INCLUDE，（S1、S2…）），表示只接收来自指定组播源S1、S2……发往组G的数据；或（G，EXCLUDE，（S1、S2…）），表示接收除了组播源S1、S2……之外的组播源发给组G的数据，即S1、S2……在接收范围之外。

（3）当组播组与组播源列表的对应关系发生了变化，在组播成员发给查询器的IGMPv3报告报文的组记录（Group Record）字段中做出相应变化。组记录有 6种类型，如表12-3所示。

表12-3 IGMPv3报告报文中的组记录类型

分类	组记录类型	含义
当前状态报告：用于对查询报文进行响应，通告自己目前的状态	IS_IN	表示组播组与源列表之间的对应方式为 INCLUDE（包括），即接收从指定源列表发往该组播组的数据
	IS_EX	表示组播组与源列表之间的对应方式为 EXCLUDE（排除），即接收从指定源列表以外的组播源发往该组播组的数据
过滤模式改变报告：用于在组播组和组播源的关系发生改变时，通告过滤模式的变化	TO_IN	表示组播组与组播源列表之间的对应方式由 EXCLUDE 转换到 INCLUDE。如果这时指定源列表为空，则表示离开该组播组，因为其中已无组播源了
	TO_EX	表示组播组与组播源列表之间的对应方式由 INCLUDE 转换到 EXCLUDE。如果这时指定源列表为空，则表示接收所有组播源发来的数据

（续表）

分类	组记录类型	含义
源列表改变报告：用于在指定组播源发生改变时，通告组播源列表的变化	ALLOW	表示在现有的基础上，还希望从某些组播源接收组播数据。如果当前对应关系为 INCLUDE，则向现有源列表中添加某些组播源；如果当前对应关系为 EXCLUDE，则从现有源列表中删除某些组播源
	BLOCK	表示在现有的基础上，不再希望从某些组播源接收组播数据。如果当前对应关系为 INCLUDE，则从现有源列表中删除某些组播源；如果当前对应关系为 EXCLUDE，则向现有源列表中添加某些组播源。它是与上面的 ALLOW 类型完全相反

在IGMPv3中一个成员报告报文可以携带多个组播组信息（而之前的IGMP版本一个成员报告只能携带一个组播组），所以在IGMPv3组播中的报文数量会大大减少。

IGMPv3没有定义专门的成员离开报文，成员离开通过特定类型的报告报文来传达。例如组225.1.1.1的成员想离开这个组，则会发送（225.1.1.1，TO_IN，（0））的报告报文，通过清空里面的指定组播源来预示要离开对应的组播组。

2．IGMPv3工作机制

在工作机制上，与IGMPv2相比，IGMPv3增加了主机对组播源的选择能力，包括特定源组加入和特定源组查询两方面。

（1）特定源组加入

IGMPv3的成员报告报文的目的地址为224.0.0.22（代表同一网段所有使能**IGMPv3**的路由器，也是一个永久组播地址 ）。通过在报告报文中携带组记录，主机在加入组播组的同时能够明确要求接收，或不接收特定组播源发出的组播数据。如图12-13所示，网络中存在S1和S2两个组播源，均向组播组G发送组播数据，但Host仅希望接收从组播源S1发往组播组G的信息。

如果主机和路由器之间运行的是IGMPv1或IGMPv2，Host加入组播组G时无法对组播源进行选择，无论其是否需要，都会同时接收到来自组播源S1和S2的数据。但如果运行的是IGMPv3，Host可以选择仅接收S1组播数据。具体有以下两种方法。

① Host发送 IGMPv3报告（G，IS_IN，（S1）），明确指定仅接收组播源S1向组播组G发送的数据。这种方法最彻底，不受后面新增组播源的影响，均只接收来自S1的数据。

② Host发送 IGMPv3报告（G，IS_EX，（S2）），明确排除不接收指定源 S2向组播组G发送的数据，这样一来就间接地预示着仅接收来自S1的组播数据。这种方法不彻底， 因为如果网络中有新增的组播要向组播组G发送数据时仍不能被排除。

（2）特定源组查询

当查询器接收到改变组播组与组播源列表的对应关系的报告时（如表12-3所示的后4种报告报文），会向组播成员发送特定源组查询报文。如果组播成员希望接收其中任意一个源的组播数据，将反馈报告报文。路由器根据反馈的组成员报告更新该组对应的源列表。

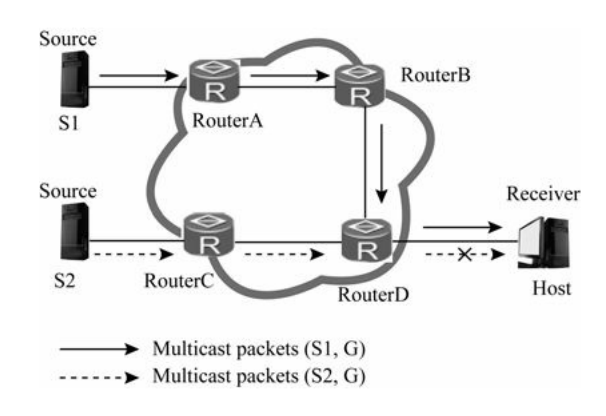


图12-13 特定源组的组播源过滤示意图

12.2.4 IGMP SSM Mapping

SSM要求路由器能了解成员主机加入组播组时所指定的组播源。如果成员主机上运行IGMPv3，可以在IGMPv3报告报文中直接指定组播源地址。但是某些情况下，用户主机只能运行IGMPv1或IGMPv2，为了使其也能够使用SSM服务，路由器上需要提供IGMP SSM Mapping功能。但 IGMP SSM Mapping不处理IGMPv3的报告报文。

SSM Mapping的机制是：通过在路由器上静态配置SSM地址的映射规则，将IGMPv1和IGMPv2报告报文中的（*，G）信息转化为对应的（S，G）信息，以提供SSM组播服务。缺省情况下，SSM组播组的组播IP地址范围为232.0.0.0～232.255.255.255。

配置了SSM Mapping规则后，当路由器收到来自成员主机的 IGMPv1或 IGMPv2报告报文时，首先检查该报文中所携带的组播组地址G，然后根据检查结果的不同分别进行处理。

（1）如果G在ASM（Any-Source Multicast，任意源组播）范围内，则只提供ASM服务。

（2）如果G在SSM组地址范围内，而路由器上又没有G对应的SSM Mapping规则，则无法提供SSM服务，丢弃该报文。

（3）如果G在SSM组地址范围内，路由器上有G对应的SSM Mapping规则，则依据规则将报告报文中所包含的（*，G）信息映射为（S，G）信息，提供SSM服务。

如图12-14所示，在SSM网络中HostA运行IGMPv3、HostB运行IGMPv2、HostC运行IGMPv1，且HostB和HostC无法升级到IGMPv3。如果要为该网段中的所有主机提供SSM服务，需要在Router上使用 IGMP SSM Mapping。

假如在Router上配置以下4个组播组（均为SSM组播IP地址）和组播源的映射关系：

- 232.0.0.0/8→10.10.1.1
- 232.1.0.0/16→10.10.2.2
- 232.1.0.0/16→10.10.3.3
- 232.1.1.0/24→10.10.4.4

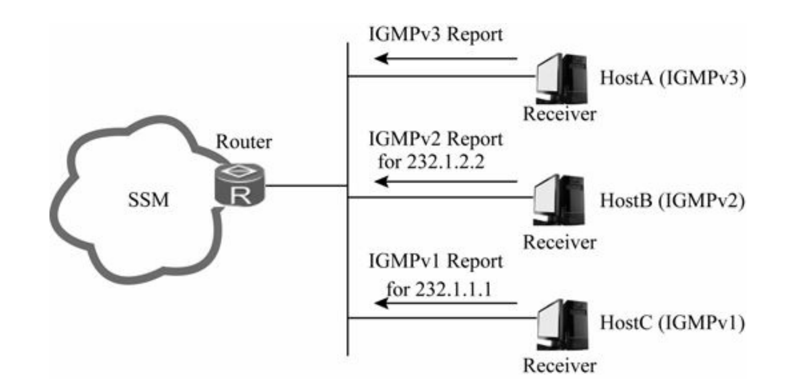


图12-14 SSM Mapping应用示例

经过映射后，Router收到HostB和HostC的成员报告报文时，首先判断报文携带的组播组IP地址是否在SSM范围内，结果发现是在SSM范围内，然后根据配置的映射规则生成如表12-4所示的组播表项。如果一个组地址映射了多个源，则生成多个（S，G）表项。在映射过程中，一个组播组地址只要能在规则中匹配到一个组播源地址，就会生 成一条相应的表项。如HostB的报告报文中携带的组播组IP地址为232.1.2.2，从

前面的组播组和组播源映射中可以看出，这样一个组播组IP地址可以分别与10.10.1.1、10.10.2.2和 10.10.3.3匹配，因为 IP 地址 232.1.2.2 可以是 232.0.0.0/8 中的一个地址，也可以是232.1.0.0/16中的一个地址（它又与10.10.2.2和10.10.3.3这两个组播源进行了映射），但不可能是232.1.1.0/24中的一个IP地址（因为第三个八位不可能一样），所以组播组地址232.1.2.2最终有3条表项。同理，组播组地址232.1.1.1最终有4条表项。

表12-4 为HostB和HostC生成的组播表项

IGMPv1/IGMPv2 报告报文中的组地址	生成的组播表项
232.1.2.2 （来自 HostB）	（10.10.1.1，232.1.2.2） （10.10.2.2，232.1.2.2） （10.10.3.3，232.1.2.2）
232.1.1.1 （来自 HostC）	（10.10.1.1，232.1.1.1） （10.10.2.2，232.1.1.1） （10.10.3.3，232.1.1.1） （10.10.4.4，232.1.1.1）

12.2.5 IGMP典型应用

IGMP 运行在成员主机和与其直连的组播路由器上，负责组播组成员关系的管理和维护。同时，为了将组播源的数据顺利转发到接收者，组播路由器之间需要运行组播路由协议PIM来建立转发路径。图12-15所示为IGMP的典型应用组网图。在实际应用中可用的方案如表12-5所示。

表12-5 IGMP的几种典型应用识方案

方案	成员主机	与成员主机相连的路由器接口	网络中的所有路由器
ASM 主机动态接入	启用 IGMPv1 或 IGMPv2 协议	启用 IGMPv1 或 IGMPv2 协议	启用 PIM-DM 或 PIM-SM 协议
SSM 主机动态接入	启用 IGMPv3 协议	启用 IGMPv3 协议	启用 PIM-SSM 协议
SSM Mapping 主机动态接入	启用 IGMPv1 或 IGMPv2 协议	启用 IGMPv3 协议，使能 SSM Mapping 功能，配置源和组映射关系	启用 PIM-SSM 协议

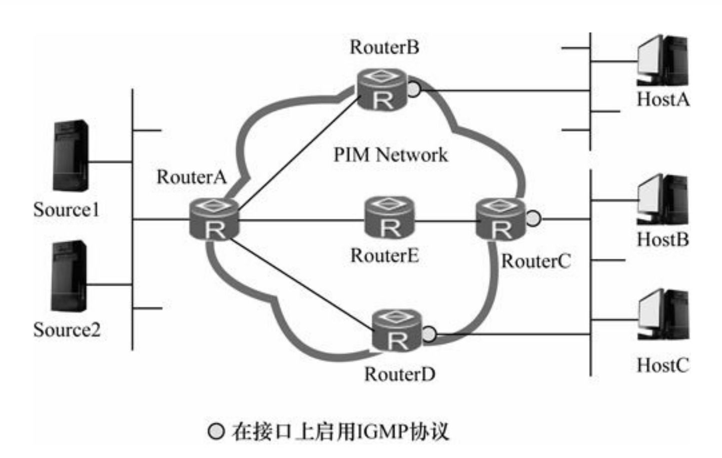


图12-15 IGMP典型应用示例

12.3 PIM基础及工作原理

PIM（协议无关组播）中的“协议无关”指的是与单播路由协议无关，即PIM不需要维护专门的单播路由信息，而是直接利用单播路由表的路由信息（注意：还有自己的组播路由表 ），对组播报文执行

RPF（Reverse Path Forwarding，逆向路径转发）检查，检查通过后创建组播路由表项，从而转发组播报文。有关RPF原理将在本章后面具体介绍。

目前在实际网络中，PIM主要有两种模式：PIM-DM（PIM-Dense Mode，PIM密集模式）、PIM-SM（PIM-Sparse Mode，PIM稀疏模式），均可用于 IPv4和 IPv6网络。由PIM路由器所组成的网络称为PIM网络。通常一个大的PIM网络可以划分为多个PIM域来管理和控制组播报文的转发，这里的域内组播协议即是指PIM域内组播协议。

在目前的PIM协议中，主要实现方式包括PIM-DM、PIM-SM（ASM模型）、PIM-SM （SSM模型）3种。SSM模型与ASM模型之间的最大差异就是是否指定了组播源，具体的区别如表12-6所示。

表12-6 PIM实现方式比较

协议	模型分类	适用场景	工作机制
PIM-DM	ASM 模型	适合规模较小、组播组成员相对比较密集的局域网	通过周期性"扩散-剪枝"维护一棵连接组播源和组成员的单向无环 SPT
PIM-SM	ASM 模型	适合网络中的组成员相对比较稀疏，分布广泛的大型网络	采用接收者主动加入的方式建立组播分发树，需要维护 RP、构建 RPT、注册组播源
	SSM 模型	适合网络中的用户预先知道组播源的位置，直接向指定的组播源请求组播数据的场景	采用 PIM-SM 的部分技术，**直接在组播源与组成员之间建立 SPT，无需维护 RP、构建 RPT、注册组播源**

下面先来了解PIM网络中的相关基本概念，然后具体介绍以上3种PIM实现方式的各功能实现原理。

12.3.1 PIM基本概念

如图12-16是一个典型的单域PIM网络，下面通过这个示例来介绍PIM的一些基本概念。

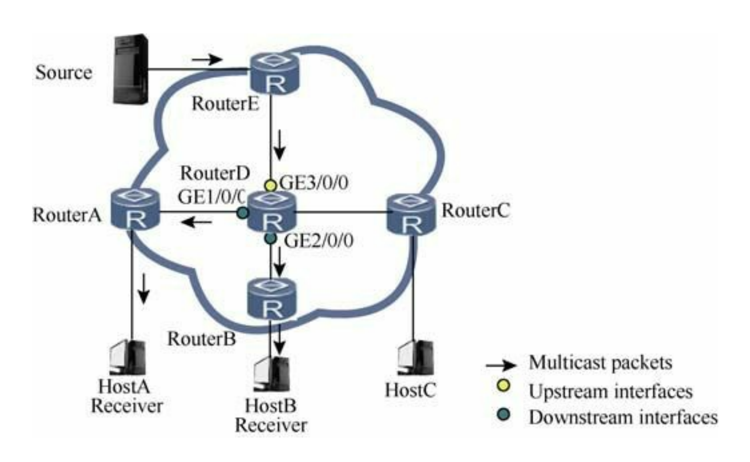


图12-16 典型单域PIM网络

1．PIM路由器

在接口上使能了PIM协议的路由器即为PIM路由器。在建立组播分发树的过程中， PIM路由器又分为以下几种。

（1）第一跳路由器：在组播转发路径上与组播源相连且负责转发该组播源发出的组播数据的PIM路由器。如图12-16中的RouterE。

（2）叶子路由器：与用户主机相连的PIM路由器，但连接的用户主机不一定为组成员，如图12-16中的RouterA、RouterB、RouterC。

（3）最后一跳路由器：在组播转发路径上与组播组成员相连，且负责向该组成员转发组播数据的PIM路由器。如图12-16中的RouterA、RouterB。

（4）中间路由器：在组播转发路径上第一跳路由器与最后一跳路由器之间的PIM路由器。如图12-16中的RouterD。

2．组播分发树

“组播分发树”是PIM网络中以组播组为单位，在PIM路由器上建立的点到多点的组播转发路径。由于组播转发路径呈现树型结构，也称为组播分发树（MDT，Multicast Distribution Tree）。

组播分发树主要包括以下两种。

（1）以组播源为根，以组播组成员为叶子的组播分发树称为SPT（Shortest Path Tree，最短路径树）。SPT 同时适用于 PIM-DM 网络和 PIM-SM 网络。如图中的RouterE→RouterD→RouterA/RouterB/RouterC就是一棵以Source为根，以HostA、HostB和HostC为叶子的SPT。

（2）以**RP**（**Rendezvous Point**，汇集点）为根，以组播组成员为叶子的组播分发树称为RPT（RP Tree，汇集点树）。RPT仅适用于PIM-SM网络。RP是通过手动配置的，具体RP的用途及工作原理将在本章12.3.3节介绍。

3．PIM路由表项

PIM路由表项即通过PIM协议建立的组播路由表项。PIM路由表项中主要用于指导转发的信息，包括组播源IP地址（是一个单播IP地址）、组播组IP地址（是一个组播IP地址）、上游接口（本地路由器上接收到组播数据的接口，如图中RouterD的GE3/0/0接口）和下游接口（将组播数据转发出去的接口，如图中RouterD的GE1/0/0和GE2/0/0接口）。

PIM网络中存在两种路由表项：（S，G）路由表项或（*，G）路由表项。S表示组播源IP地址，G表示组播组IP地址，*表示任意组播源。其中，（S，G）路由表项中明确指定了组播源S的位置，主要用于在PIM路由器上建立SPT（最短路径树），同时适用于PIM-DM和PIM-SM网络；而（*，G）路由表项中代表不知道组播源位置，只知道组播组IP地址，主要用于在PIM路由器上建立RPT（汇集点树），仅适用于PIM-SM网络。

PIM路由器上可能同时存在以上两种路由表项。当收到源地址为S，组地址为G的组播报文，且通过RPF（逆向路径转发）检查的情况下，按照如下的规则转发。

（1）如果存在（S，G）路由表项，则由（S，G）路由表项指导报文转发。

（2）如果不存在（S，G）路由表项，只存在（*，G）路由表项，则先依照（*，G）路由表项创建（S，G）路由表项，再由（S，G）路由表项指导报文转发。

12.3.2 PIM-DM基本工作原理

PIM-DM（PIM 密集模式）使用“推”（Push）模式转发组播报文，一般应用于组播组成员规模相对较小、相对密集的网络。在实现过程中，它会假设网络中的组成员分布非常稠密，每个网段都可能存在组成员。当有活跃的组播源出现时，PIM-DM会将组播源发来的组播报文扩散到整个网络的PIM路由器上，再裁剪掉不存在组播报文转发的分支。

PIM-DM就这样通过周期性地进行“扩散（Flooding）——剪枝（Prune）”过程来构建并维护一棵连接组播源和组成员的单向无环SPT（Source Specific Shortest Path Tree，源指定最短路径树）。如果在下一次“扩散——剪枝”进行前，被裁剪掉的分支由于其叶子路由器上有新的组成员加入而希望提前恢复转发状态，也可通过嫁接（Graft）机制主动恢复其对组播报文的转发。

综上所述，PIM-DM的关键工作机制包括邻居发现、扩散、剪枝、嫁接、断言和状态刷新。其中，扩散、剪枝、嫁接是构建SPT的主要方法。下面分别予以介绍。

1．邻居发现（Neighbor Discovery）

在PIM路由器每个使能了PIM协议的接口上都会对外发送Hello报文。封装Hello报文的组播报文的目的地址是224.0.0.13（代表同一网段中所有**PIM**路由器，是一个永久组播地址 ）、源地址为接口的IP地址、TTL数值为1。

Hello报文的作用：发现PIM邻居、协调各项PIM协议报文参数，并维持邻居关系。

在发现 PIM 邻居的过程中，同一网段中的 PIM 路由器都必须接收目的地址为224.0.0.13的组播报文。这样直接相连的PIM路由器之间通过交互Hello报文后就可以彼此知道自己的邻居信息，建立邻居关系。只有邻居关系建立成功后，PIM路由器之间才能相互接收PIM协议报文，从而创建组播路由表项。

Hello报文中携带多项PIM协议报文参数，主要用于PIM邻居之间PIM协议报文的控制，协调各项PIM协议报文参数。具体参数包括以下几种。

（1）DR_Priority：表示各路由器接口竞选 DR（指定路由器）的优先级，优先级越高越容易获胜，担当IGMPv1的查询器（注意，如果是运行IGMPv2或IGMPv3则采用专门的查询器选举机制），参见本章前面12.2.2节。

（2）Holdtime：表示保持邻居为可达状态的超时时间，超过这个时间没收到邻居发来的Hello报文即认为该邻居不可达。这与RIP、OSPF等动态路由协议中的Hello报文是一样的。

（3）LAN_Delay：表示共享网段内传输 Prune（剪枝）报文的延迟时间，超过这个时间，这个报文将被丢弃。

（4）Neighbor-Tracking：表示邻居跟踪功能。

（5）Override-Interval：表示 Hello 报文中携带的否决剪枝的时间间隔。当超过这个时间后原来的剪枝状态就要被中止，恢复对应出接口的组播转发功能。

2．维持邻居关系

PIM路由器之间周期性地发送Hello报文。如果Holdtime超时还没有收到该PIM邻居发出的新的Hello报文，则认为该邻居不可达，将其从邻居列表中清除。

PIM邻居的变化将导致网络中组播拓扑的变化。如果组播分发树上的某上游邻居或 下游邻居不可达，将导致组播路由重新收敛，组播分发树迁移。

3．扩散

当PIM-DM网络中出现活跃的组播源之后，组播源发送的组播报文将在全网内扩散（Flooding）。“扩散”的目的其实就是为了下一步的“剪枝”和“断言”操作。当PIM路由器接收到组播报文，并根据单播路由表进行RPF检查通过后，就会在该路由器上创建（S，G）表项。在PIM路由器的下游接口列表中包括了除上游接口之外，与所有PIM邻居相连的接口，到达的组播报文将从各个下游接口转发出去。最后组播报文扩散到达叶子路由器，此时会出现以下两种情况。

（1）如果与该叶子路由器相连用户网段上存在组成员，则将与该网段相连的接口加入（S，G）表项的下游接口列表中，后续的组播报文会向组成员转发。

（2）如果与该叶子路由器相连用户网段上不存在组成员，且不需要向其下游PIM邻居转发组播报文，则执行“剪枝”机制，从组播路径中去掉这部分路径。具体将在下面介绍。

说明

有时组播报文扩散到一个连着多台PIM路由器的共享网段时，会出现这种情况： 这些 PIM 路由器上进行的 RPF 检查都能通过，从而有多份相同报文转发到这个网段。此时，需要执行“断言”机制，保证只有一

个PIM路由器向该网段转发组播报文。具体将在下面介绍。

如图12-17所示，在PIM-DM网络中，RouterA、RouterB和RouterC之间通过发送Hello报文建立了PIM邻居关系。HostA通过RouterA与HostA之间运行的IGMP协议加入了组播组 G，HostB 没有加入任何组播组。下面看一下本示例中扩散的具体过程，从中也反映了扩散的目的之一——“剪枝”。

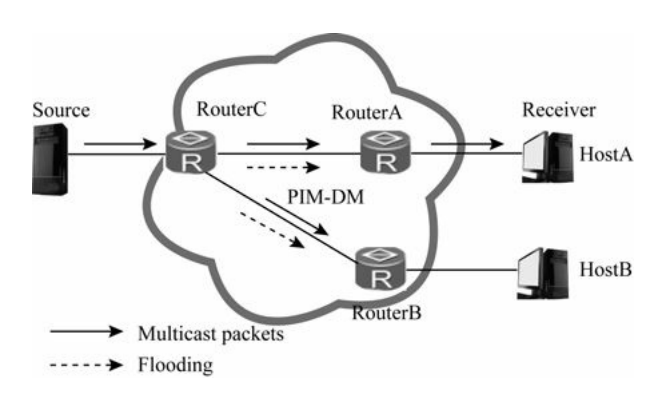


图12-17 扩散示意图

（1）组播源S开始向组播组G发送组播报文。

（2）RouterC接收到源发送的组播报文后根据单播路由表进行RPF检查。RPF检查通过后创建（S，G）表项，下游接口列表包括与RouterA和RouterB相连的接口，后续到达的报文向RouterA和RouterB转发。

（3）RouterA接收来自RouterC的组播报文，通过RPF成功检查后在本地创建对应（S，G）表项，并在下游接口列表添加与组成员HostA相连的接口，后续到达的报文向HostA转发。

（4）RouterB接收来自RouterC的组播报文，由于与RouterB相连下游网段不存在组 成员和PIM邻居，所以执行剪枝操作，不会发送组播数据到HostB上。

4．剪枝（Prune）

通过上面介绍的“扩散”特性了解了“剪枝”的目的，本节要具体介绍“剪枝”的原理。

当 PIM 路由器接收到组播报文后，通过 RPF 检查，但是下游网段没有组播报文需求时，PIM路由器会向上游发送剪枝报文，通知上游路由器禁止相应下游接口的转发，将其从（S，G）表项的下游接口列表中删除。剪枝操作由叶子路由器发起，逐跳向上，最终组播转发路径上只存在与组成员相连的分支。

路由器为被裁剪的下游接口启动一个剪枝定时器，定时器超时后接口恢复转发。这时，组播报文又会重新在全网范围内扩散，新加入的组成员可以接收到组播报文。随后，下游不存在组成员的叶子路由器再次将向上发起剪枝操作。通过这种周期性地扩散——剪枝，PIM-DM周期性地刷新SPT。当下游接口被剪枝后会执行以下操作。

（1）如果下游叶子路由器有组成员加入，并且希望在下次“扩散——剪枝”前就恢复组播报文转发，则执行“嫁接”机制。具体将在下面介绍。

（2）如果下游叶子路由器一直没有组成员加入，希望该接口保持抑制转发状态，则执行“状态刷新机制”。 具体也将在下面介绍。

如图12-18所示，RouterB上未连接组成员。在这种情况下，RouterB会向上游发起剪枝请求。具体过程如下。

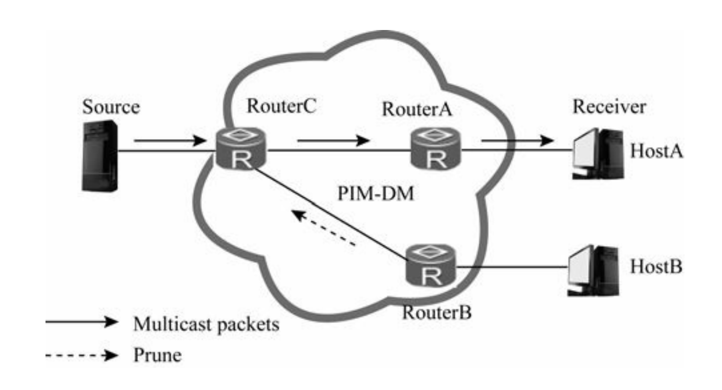


图12-18 剪枝示意图

（1）RouterB向上游RouteC发送Prune报文，通知RouterC不用再转发数据到该下游网段。

（2）RouterC收到Prune报文后，停止该下游接口（也就是与RouterB相连的出接口）转发，将该下游接口从（S，G）表项中删除，后续到达的报文只向RouterA转发。

5．嫁接（Graft）

PIM-DM通过“嫁接机制”可使有新组成员加入的网段快速得到组播报文。叶子路由器通过IGMP了解到与其相连的用户网段上，组播组G有新的组成员加入。随后叶子路由器会向上游发送Graft报文，请求上游路由器恢复相应出接口转发，将其添加在（S， G）表项下游接口列表中。

嫁接过程从叶子路由器开始，到有组播报文到达的路由器结束。在如图12-19所示的示例中的具体嫁接过程如下。

（1）RouterB 希望立即恢复对 HostB 组播报文的转发，于是向上游路由器 RouterC 发送Graft报文，请求恢复相应出接口转发组播报文。

（2）RouterC收到Graft报文后，恢复与RouterB相连的出接口转发，将该接口添加到（S，G）表项中的下游接口列表中，这样后续到达的报文向RouterB转发，直达HostB。

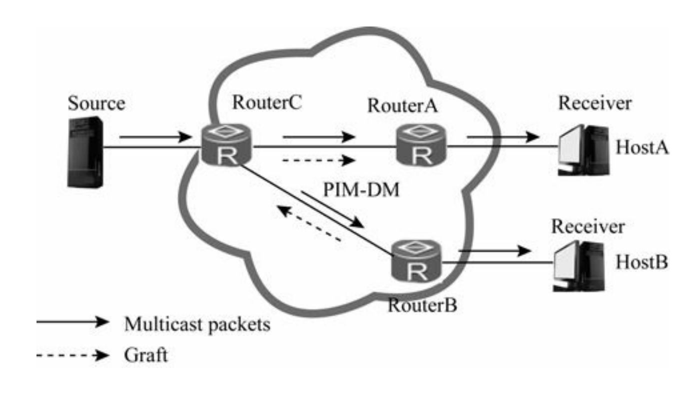


图12-19 嫁接示意图

6．状态刷新（State Refresh）

在 PIM-DM 网络中，为了避免被裁剪的接口因为“剪枝定时器”超时而恢复转发，离组播源最近的第一跳路由器会周期性地触发State Refresh报文在全网内扩散。收到状态刷新（State Refresh）报文的PIM路由器会刷新剪枝定时器的状态，其目的就是查找原来被剪枝的路径上是否有组播成员要加入，要恢复转发状态，是一种被动中止剪枝状态的操作。在原来被裁剪接口的下游叶子路由器上，如果有新的组成员加入，则立即中止剪枝状态，对应路径的组播转发；如果仍没有组成员加入，则该接口将一直处于抑制转发状

态。

7．断言（Assert）

当一个网段内有多个相连的 PIM 路由器通过 RPF 检查后向该网段转发相同的组播报文时，则需要通过“断言机制”来保证只有一个PIM路由器向该网段转发组播报文，以保证组成员不接收多份相同的组报文。

这个“断言机制”是在PIM路由器接收到邻居路由器发送的相同组播报文后，以组播的方式向本网段的所有PIM路由器发送Assert报文，目的地址为224.0.0.13（代表所有PIM路由器）。其他PIM路由器在接收到Assert报文后，将自身参数与对方报文中携带的参数做比较，进行Assert竞选。竞选规则如下。

（1）单播路由协议优先级较高者获胜。

（2）如果优先级相同，则到组播源的路径开销较小者获胜。

（3）如果以上都相同，则下游接口IP地址最大者获胜。

根据Assert竞选结果，路由器将执行不同的操作。

（1）获胜一方的下游接口称为Assert Winner，将负责后续对该网段组播报文的转发。

（2）失败一方的下游接口称为Assert Loser，后续不会对该网段转发组播报文，PIM路由器也会将其从（S，G）表项下游接口列表中删除。

Assert竞选结束后，该网段上只存在一个下游接口，只传输一份组播报文。所有Assert Loser可以周期性地恢复组播报文转发，从而引发周期性的Assert竞选。

如图12-20所示，RouterB和RouterC均通过了RPF检查，创建了（S，G）表项，并且两者的下游接口连接在同一网段，RouterB 和 RouterC 都向该网段发送组播报文。 具体断言过程如下。

（1）RouterB和RouterC从各自上游接口接收到RouterA发来的组播报文，RPF检查都失败，报文被丢弃。这时，RouterB和RouterC就会分别向该网段发送Assert报文。

（2）RouterB在收到RouterC发来的Assert报文后，将自身的路由信息与Assert报文中携带的路由信息进行比较，由于RouterB自身到组播源的开销较小而获胜。于是后续组播报文仍然向该网段转发，RouterC在接收到组播报文后仍然由于RPF检查失败而丢弃。

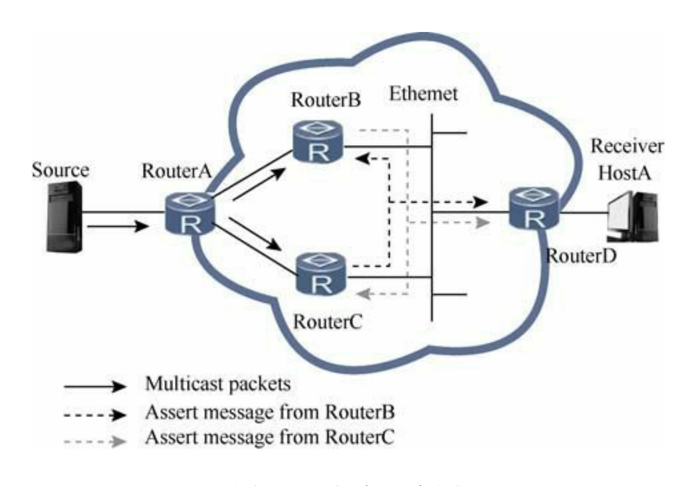


图12-20 断言示意图

（3）同样，RouterC在收到RouterB发来的Assert报文，也将自身的路由信息与报文中携带的路由信息进行比较，由于RouterC自身到组播源的开销较大而落败。于是禁止相应下游接口向该网段转发组播报文，将其从（S，G）表项的下游接口列表中删除。

12.3.3 PIM-SM（ASM模型）工作原理

PIM-SM适用于ASM和SSM两种模型。在ASM模型中，它使用"拉（Pull）模式"转发组播报文，一般应用于组播组成员规模相对较大、相对稀疏的网络。其基本工作机制如下。

（1）在网络中维护一台 RP，可以为随时出现的组成员或组播源服务。网络中所有PIM路由器都知道RP的位置。

（2）当网络中出现组成员（用户主机通过IGMP加入某组播组G）时，最后一跳路由器向RP发送Join报文，逐跳创建（*，G）表项，生成一棵以RP为根的RPT。

（3）当网络中出现活跃的组播源时（信源向某组播组 G 发送第一个组播数据时），第一跳路由器将组播数据封装在Register报文中单播发往RP，在RP上创建（S，G）表项，注册源信息。

在ASM模型中，PIM-SM的关键机制包括邻居发现、DR竞选、RP发现、RPT构建、组播源注册、SPT切换、剪枝、断言；同时也可通过配置BSR（Bootstrap Router，自举路由器）管理域来实现单个 PIM-SM 域的精细化管理。其中"邻居发现""断言机制"与上节PIM-DM中介绍的"邻居发现"和"断言机制"是完全一样的，参见即可。

1．DR竞选

在组播源或组成员所在的网段，通常同时连接着多台 PIM 路由器。这些 PIM 路由器之间通过交互Hello报文成为PIM邻居，Hello报文中携带DR优先级和该网段接口地 址。PIM 路由器将自身条件与对方报文中携带的信息进行比较，选举出唯一的 DR（注意：每个网段要选举一个**DR**，并不是整个组播网络中只能有一台**DR** ）来负责源端或组成员端组播报文的收发。竞选规则如下。

（1）DR优先级较高者获胜（在网段中所有PIM路由器都支持DR优先级的情况下）。

（2）如果DR优先级相同，或该网段存在至少一台PIM路由器不支持在Hello报文中携带DR优先级，则IP地址较大者获胜。

（3）如果当前DR出现故障，将导致PIM邻居关系超时，其他PIM邻居之间会触发新一轮的DR竞选。

在ASM模型中DR主要作用如下。

（1）在连接组播源的共享网段，由DR负责向RP发送Register注册（组播源注册）报文。与组播源相连的DR称为源端DR。

（2）在连接组成员的共享网段，由DR负责向RP发送Join加入（组成员加入）报文。与组成员相连的DR称为组成员端DR。

2．RP发现

RP为网络中一台重要的PIM路由器，用于处理组播源DR注册信息及组成员加入请求，网络中的所有PIM路由器都必须知道RP的地址，类似于一个供求信息的汇聚中心。

一个RP可以同时为多个组播组服务，但一个组播组只能对应一个**RP** 。目前可以通过以下方式配置RP。

（1）静态RP：在网络中的所有PIM路由器上配置相同的RP地址，静态指定RP的位置。

（2）动态RP：在PIM域内选择几台PIM路由器，配置C-RP（Candidate-RP，候选RP）来动态竞选出RP。不过此时，还需要通过配置C-BSR（Candidate-BSR，候选BSR）选举出BSR，来收集C-RP的通告信息，向PIM-SM域内的所有PIM路由器发布。

说明

BSR（自举路由器）是PIM-SM网络里的管理核心，负责收集网络中C-RP（Candidate-RP，候选RP）发来的宣告信息（Advertisement message），然后将为每个组播组选择部分C-RP信息组成RP-Set（即组播组和RP的映射数据库），并以BSR消息（BSR message）发布到整个PIM-SM网络，从而使网络内的所有路由器（包括DR）都知道RP的位置。

BSR的选举过程中，初始时每个C-BSR都认为自己是BSR，向全网发送Bootstrap报文。Bootstrap报文中携带C-BSR地址、C-BSR的优先级。每一台PIM路由器都收到所有C-BSR发出的Bootstrap报文，通过比较这些C-BSR信息，竞选产生BSR。BSR的竞选规则如下。

（1）C-BSR优先级较高者获胜（优先级数值越大优先级越高）。

（2）如果优先级相同，IP地址较大的C-BSR获胜。

C-RP竞选的具体过程。

（1）C-RP向BSR发送Advertisement报文，报文中携带C-RP地址、服务的组范围和C-RP优先级。

（2）BSR收到这些Advertisement报文后，将这些信息汇总为RP-Set（RP集），封装在Bootstrap报文中，发布给全网的每一台PIM-SM路由器。

（3）各PIM路由器收到Bootstrap报文后，使用相同的规则进行计算和比较，从多 个针对特定组的C-RP中竞选出该组RP。规则如下。

① C-RP接口地址掩码最长者获胜。

② C-RP优先级较高者获胜（优先级数值越大优先级越低）。

③ 如果优先级相同，则执行Hash函数，计算结果较大的C-RP获胜。

④ 如果以上都相同，则C-RP地址较大者获胜。

由于所有PIM路由器使用相同的RP-Set和竞选规则，所以得到的组播组与RP之间的对应关系也相同。PIM路由器将“组播组——RP”对应关系保存下来，指导后续的组播操作。

3．RPT构建

PIM-SM RPT 是一棵以 RP 为根，以存在组成员关系的 PIM 路由器为叶子的组播分发树，如图12-21所示。当网络中出现组成员（用户主机通过IGMP加入某组播组G）时，组成员端DR向RP发送Join报文，在通向RP的路径上逐跳创建（*，G）表项，生成一棵以RP为根的RPT。

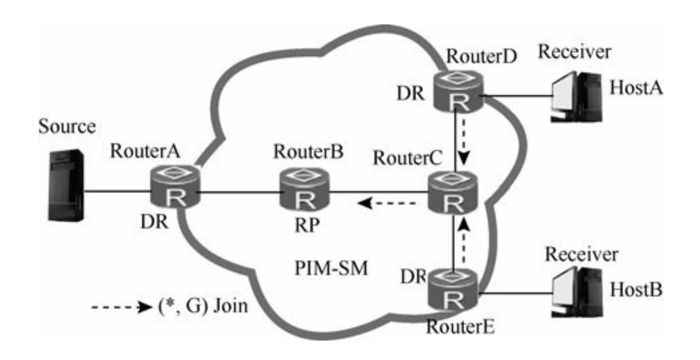


图12-21 RPT示例

在RPT构建过程中，PIM路由器在收发 Join 报文时，都会进行 RPF 检查。接收者DR首先执行RPF检查：查找到达RP的单播路由，单播路由的出接口为上游接口，下一跳为RPF邻居。然后，接收者DR向该RPF邻居发送Join报文。RPF邻居接收到Join报文后，执行RPF检查，如果检查通过，继续向上游发送。Join报文逐跳上送，直至到达RP。

4．组播源注册

组播源注册是也是在 RP 上进行的，但注册信息是通过源端 DR 传递到 RP 的。在PIM-SM 网络中，任何一个新出现的组播源都必须首先在 RP 处注册，然后才能将组播报文传输到组成员。具体过程如下。

（1）组播源将组播报文发给源端DR。

（2）源端DR接收到组播报文后，将其封装在Register报文中，发送给RP。

（3）RP接收到Register报文后，将其解封装，并根据报文中的信息建立对应（S， G）表项，然后将组播数据沿RPT发送到达组成员。

5．SPT切换

在PIM-SM网络中，一个组播组只对应一个RP，只构建一棵**RPT** 。在未进行SPT切换的情况下，所有发往该组的组播报文都必须先封装在注册报文中发往RP，RP解封装后，再沿RPT分发。但这样会出现一个问题，那就是因为RP是所有组播报文必经的中转站，当组播报文速率逐渐较大时会对RP形成巨大的负担。为了解决此问题，PIM-SM允许RP或组成员端DR通过触发SPT切换来减轻RP的负担。

RP触发SPT切换的原理：在RP收到源端DR的注册报文后，将封装在Register报文中的组播报文直接沿**RPT**转发给组成员 （不进行解封），同时RP会向源端DR逐跳发送Join报文。发送过程中在PIM 路由器创建（S，G）表项，从而建立了RP到源的 SPT。SPT树建立成功后，源端DR直接将组成员加入的组播报文转发到RP。最终使源端DR和RP免除频繁的封装与解封装。

如图12-22所示，组成员端DR周期性检测组播报文的转发速率，一旦发现（S，G）报文的转发速率超过阈值，则触发以下SPT切换。

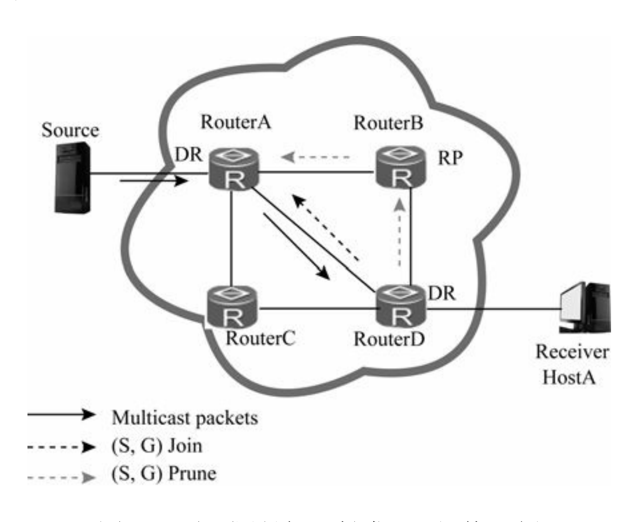


图12-22 组成员端DR触发SPT切换示例

（1）组成员端DR（如RouterD）逐跳向源端 DR 逐跳发送 Join 报文并创建（S，G）表项，建立源端DR 到组成员DR的SPT。

（2）SPT建立后，组成员端DR会沿着 RPT 逐跳向 RP 发送剪枝报文，删除（S，G）表项中相应的下游接口。剪枝结束后，RP不再沿RPT转发组播报文到组成员端。

如果SPT不经过RP，RP会继续向源端DR逐跳发送剪枝报文，删除（S， G）表项中相应的下游接口。剪枝结束 后，源端DR不再沿“源端DR-RP”的SPT转发组播报文到RP。

缺省情况下，设备一般未设置组播报文转发速率的阈值，RP或者组成员端DR在接收到第一份组播报文时都会触发各自的SPT切换。

6．BSR管理域

为了实现网络管理精细化，可以选择将一个PIM-SM网络划分为多个BSR管理域和一个Global（全局）域。这样一方面可以有效地分担单一BSR的管理压力，另一方面可以使用私有组地址为特定区域的用户提供专门服务。

每个BSR管理域中维护一个BSR，为某一特定地址范围的组播组服务。Global域中维护一个BSR，为所有剩余的组播组服务。

BSR管理域是针对特定地址范围的组播组的管理区域，属于此范围的组播报文只能在本管理域内传播，无法通过BSR管理域边界。图12-23所示包括了BSR1和BSR2两个管理域。对于有相同组地址的不同管理域，各BSR管理域所包含的PIM路由器互不相同，同一**PIM**路由器不能从属于多个**BSR**管理域 。各BSR管理域在地域上相互独立，且相互隔离。Global域包含PIM-SM网络内的全部PIM路由器。不属于任意BSR管理域的组 播报文，可以在整个PIM网络范围内传播。

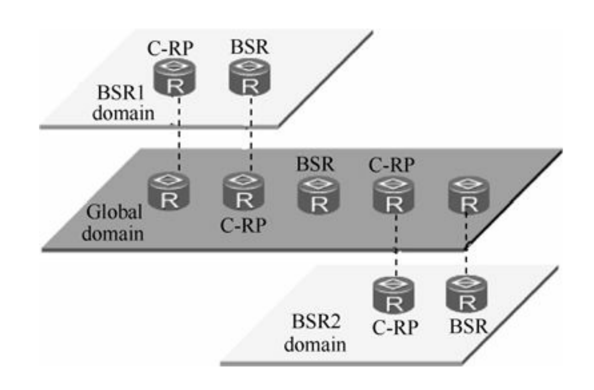


图12-23 BSR管理域示意图

如果从组播组地址范围来看，每个BSR管理域为特定地址范围的组播组提供服务，不同的 BSR 管理域服务的组播组地址范围可以重叠。但每个组播组地址只在本BSR管理域内有效，相当于私有组地址。如图12-24所示，BSR1域和BSR3域对应的组播组地址范围出现重叠。

不属于任何BSR管理域的组播组，一律 属于 Global 域的服务范围。图 12-24 中的Global 域组地址范围是除 G1、G2 之外的G-G1-G2组播地址。

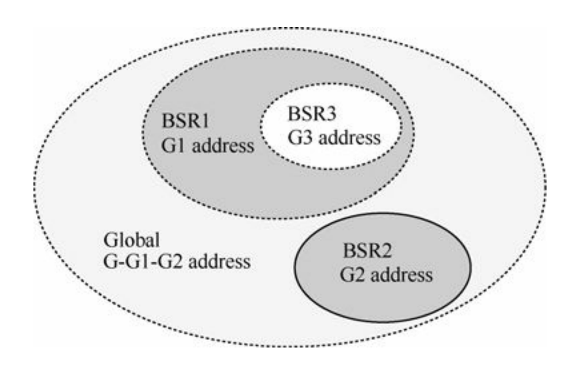


图12-24 BSR管理域的地址范围示意图

Global域和每个BSR管理域都包含针对自己域的C-RP和BSR设备，这些设备在行使相应功能时仅在本域内有效。即BSR机制和 RP 竞选在各管理域之间是隔离的。每个BSR 管理域都有自己的边界，该管理域的组播信息（C-RP宣告报文、BSR自举报文等） 不能跨越域传播。但 Global 域的组播信息可以在整个 Global 域内传递，可以穿越任意BSR管理域。

12.3.4 PIM-SM（SSM模型）工作原理

SSM模型是借助PIM-SM的部分技术和IGMPv3/MLDv2来实现的，无需维护**RP**、无需构建**RPT**、无需注册组播源 ，可以直接在源与组成员之间建立SPT。

SSM的特点是网络用户能够预先知道组播源的具体位置，因此用户在加入组播组时可以明确指定从哪

些源接收信息。组成员端DR了解到用户主机的需求后，直接向源端DR发送Join报文。Join报文逐跳向上传输，在源与组成员之间建立SPT。

在SSM模型中，PIM-SM的关键机制包括邻居发现、DR竞选、构建SPT。其中“邻居发现”机制与12.3.2节介绍的PIM-DM邻居发现机制一样，而“DR竞选”机制与12.3.3节介绍的PIM-SM（ASM模型）的“DR竞选”机制一样，分别参见即可。下面仅介绍其SPT构建原理。

下面以图12-25为例介绍PIM-SM（SSM模型）的SPT构建原理。具体过程如下。

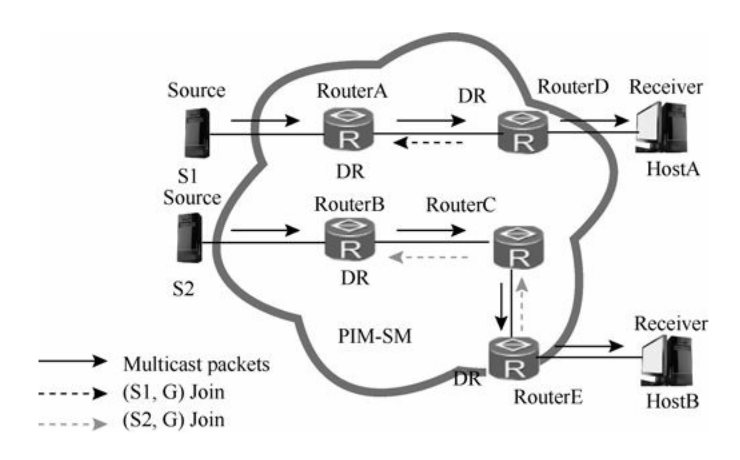


图12-25 SPT构建示例

（1）担当组成员端DR的RouterD、RouterE借助IGMPv3/MLDv2协议了解到用户主机有到相同组播组不同组播源的组播需要，于是分别逐跳向源方向（SSM模型中组播源是已知的）发送Join报文。

（2）沿途各PIM路由器通过提取Join报文中的相关信息分别创建（S1，G）、（S2， G）表项，最终就形成了从源S1到组成员HostA、源S2到组成员HostB的SPT。

（3）SPT建立后，源端就会将组播报文沿着SPT分发给组成员。

12.3.5 单自治域PIM-SM应用

大多数组播应用还是在单个AS、单个PIM域环境下的，所以在此仅介绍这种环境下的典型应用。如图12-26所示，是一个比较大型的组播网络，该网络中已经部署了完备的IGP路由，且任意网段路由可达。网络中的组成员分布相对比较稀疏，要求网络中的用户主机能够按需接收视频点播信息，并在一定程度上节约网络的带宽。

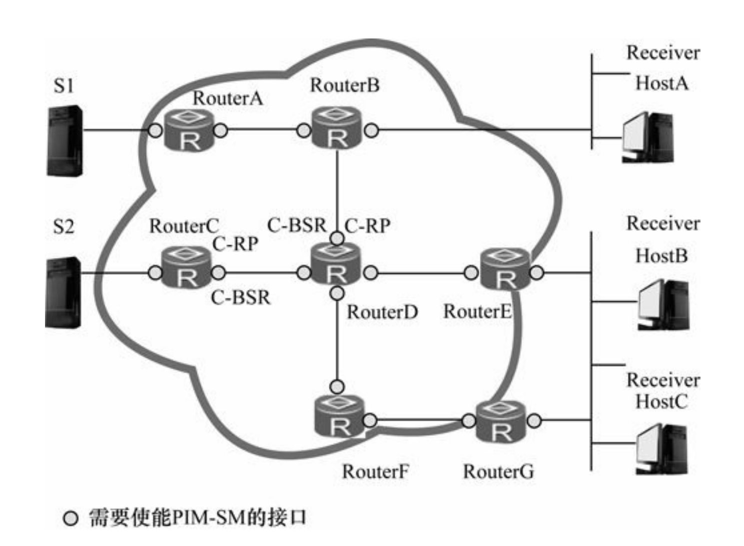


图12-26 单自治域典型PIM-SM应用示例

本示例中，HostA、HostB和HostC为三个末梢网络中的信息接收者，通过组播方式接收视频点播信息，整个PIM域采用PIM-SM方式。RouterA与组播源S1相连，RouterC与组播源S2相连；RouterB连接HostA，RouterE和RouterG连接HostB和HostC。在所有路由器接口上启用PIM-SM协议。

因为网络中的组播源分布比较密集，则可以选择与组播源比较近的核心设备作为C-RP。将RouterC和RouterD的接口配置为C-BSR和C-RP，动态竞选出为PIM-SM网络服务的BSR和RP。在RouterB与HostA之间，RouterE、RouterG与HostB、HostC之间均运行IGMP协议。

说明

RP的部署方式可以基于以下原则（避免在一个PIM域中不同路由器上分别使用静态RP和动态RP，以防止RP信息不一致）。

（1）中小型网络：建议选择静态RP方式，对设备要求低，也比较稳定。

如果网络中只有一个组播源，建议选择直连组播源的设备做为静态RP，这样可以省略组播源DR向RP注册的过程。采用静态RP方式要确保域内所有路由器（包括RP本身）的RP信息以及服务的组播组范围全网一致。

（2）大型网络：可以采用动态RP方式，可靠性高，可维护性强。

如果网络中存在多个组播源，且分布密集，建议选择与组播源比较近的核心设备作为C-RP；如果网络中存在多个用户，且分布密集，建议选择与用户比较近的核心设备作为C-RP。

为路由器接口配置IGMP协议时，请确保接口参数配置的一致性，即遵循如下原则：连接在同一网段的所有路由器必须运行相同的IGMP版本（推荐使用IGMPv2），且各接口参数（如查询定时器、组成员关系保持时间等）必须相同。如果IGMP版本或各参数不相同，会导致不同路由器上IGMP组成员关系不一致。

部署完上述网络后，HostA和HostB根据需要向RP发送Join消息，组播源的信息能够到达接收者。建议在网络边缘配置接口静态加入用户所请求的组播组，可以提高用户收看频道的稳定性。

12.4 MSDP基础及工作原理

在不同 PIM-SM 域间的 RP 信息是隔离的，缺省情况下，组播源只向本域内的 RP注册，用户主机只向本域内的RP发起加入。将一个大的PIM-SM网络划分为多个PIM-SM域后，如何实现PIM-SM域间组播，使

本PIM-SM域内的用户主机能够接收到其他域内组播源发出的组播数据就成了现实的问题。这就是 MSDP 诞生的背景。它可使不同PIM-SM域的RP之间能够互相通信，发现并共享其他PIM-SM域内的组播源信息。但目前MSDP只支持在IPv4网络部署，且仅对PIM-SM（ASM模型）有意义，因为在SSM模型中无MSDP所需的RP配置。

12.4.1 MSDP对等体概述

MSDP要互连不同PIM-SM域RP的过程就是需要在不同域之间建立对等体（peer），主要用于不同ISP网络间。通常，ISP并不希望借助其他ISP的RP来向自己的用户提供服务。这一方面是出于安全性考虑，另一方面如果其他ISP的RP发生故障导致业务中断，用户投诉的却是自己的服务。借助MSDP，每个ISP可以实现依靠自己的RP来向Internet转发和接收组播数据。

使用MSDP实现跨域组播的首要任务是建立MSDP对等体。通过配置MSDP对等体使各个PIM-SM域的RP之间建立MSDP对等体关系，各MSDP对等体之间彼此首尾相连，形成一张“MSDP连通图”，连接各PIM-SM域RP。表12-7列出了在RP上可以创建的MSDP对等体类型（参见图12-27）。

表12-7 在RP上创建的MSDP对等体类型

MSDP 对等体分类	位置	功能
源端 MSDP 对等体	离组播源（Source）最近的MSDP 对等体（通常也就是源端 RP，如 RP1）	源端 RP 创建 SA 消息并发送给远端 MSDP 对等体，通告在本 RP 上注册的组播源信息 源端 MSDP 对等体必须配置在 RP 上，否则将无法向外发布组播源信息
接收者端 MSDP 对等体	离接收者（Receiver）最近的 MSDP 对等体（如 RP3）	接收者端 MSDP 对等体在收到 SA 消息后，根据该消息中所包含的组播源信息，跨域加入以该组播源为根的 SPT；当来自该组播源的组播数据到达后，再沿 RPT 向本地接收者转发。 接收者端 MSDP 对等体也必须配置在 RP 上，否则无法接收到其他域的组播源信息
中间 MSDP 对等体	拥有多个远端 MSDP 对等体的 MSDP 对等体（如 RP2）	中间 MSDP 对等体把从一个远端 MSDP 对等体收到的 SA 消息转发给其他远端 MSDP 对等体，其作用相当于传输组播源信息的中转站

（续表）

说明

但MSDP 对等体并不是只能配置在RP上，MSDP对等体可以创建在任意的PIM路由器上，在不同角色的PIM 路由器上所创建的MSDP对等体的功能有所不同。在普通的非RP PIM路由器上也可创建MSDP对等体。如在图12-27中的RouterA和RouterB上，其作用仅限于将收到的 SA 消息转发出去。但为了保证网络中所有RP都能参与源信息共享，且尽量缩小“MSDP连通图”的规模，推荐的配置方案：在且仅在网络中所有RP上配置MSDP对等体。

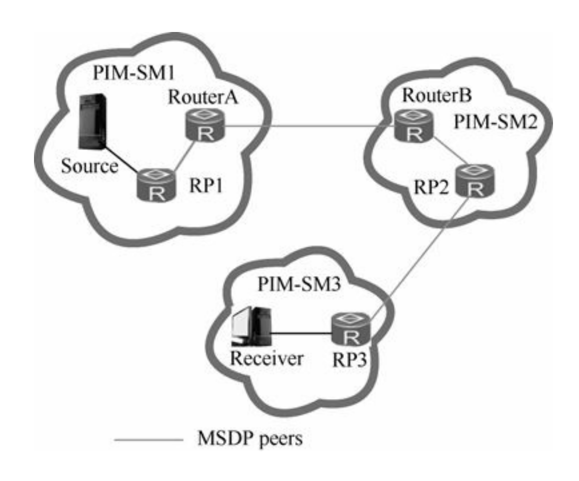


图12-27 MSDP对等体示例

12.4.2 MSDP对等体建立流程

MSDP 对等体通过 TCP 传输层协议连接建立，使用的端口为TCP 639。两台设备使能MSDP并互相指定对方为MSDP对等体后，两端先比较IP地址，IP地址较小的一端会启动连接重试定时器（ConnectRetry timer），并主动发起TCP连接。IP地址较大的一端负责确认是否有TCP连接在端口639建立。TCP连接建立后，MSDP对等体关系就建立了，对等体之间通过KeepAlive（保持活跃）消息维持连接关系。

下面以图12-28中的RouterA和RouterB之间MSDP对等体的建立为例介绍MSDP对等体建立的流程。具体如下。

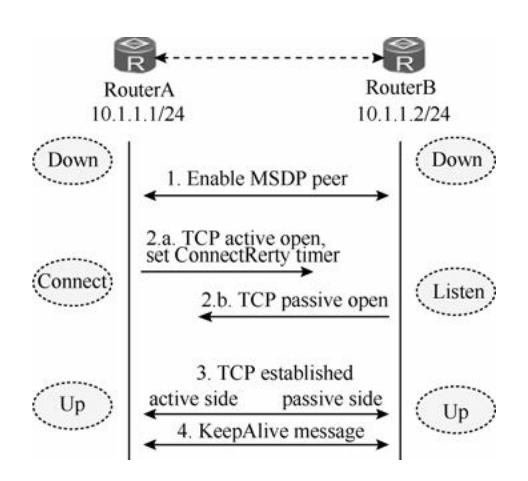


图12-28 MSDP对等体建立流程示例

（1）起始状态下，两台路由器的MSDP会话状态都是Down。

（2）在两端使能MSDP并互相指定对方为MSDP对等体后，两端比较建立连接使用的IP地址：由于RouterA的IP地址较小，所以进入Connect（连接）状态，向RouterB发起连接，并启动 ConnectRetry 定时器。该定时器用于定义连接重试的周期。RouterB的IP地址较大，此时进入Listen（监听）状态，等待对端的TCP连接。

（3）TCP连接建立成功后，两端的MSDP会话均进入Up状态，代表着MSDP对等体建立成功。随后各自向对方发送KeepAlive消息，通知对方保持MSDP连接状态。

说明

在MSDP对等体建立TCP连接过程中可以进行加密认证，以保证MSDP对等体 建立的安全性。配置了认证功能后，MSDP对等体两端必须都使用相同的加密算法和密码，才能正常建立TCP连接。MSDP支持MD5和Keychain两种加密方式，二者在使用时互斥，MSDP对等体之间只能选择一种方式加密。

12.4.3 基于MSDP的Anycast RP

前面说了，MSDP是用于PIM-SM域间的组播网络连接，但是MSDP在PIM-SM域内也有一种特殊的应用，那就是它的一种特殊RP——“任播RP”（Anycast RP）功能。它用来解决传统 PIM-SM 域中每个组播组只能映射到一个 RP，当网络负载较大或者流量过于集中时可能导致的压力过大、RP失效后收敛较慢、组播转发路径非最优等问题。

通过MSDP的Anycast RP功能，可以在同一PIM-SM域内配置多个具有相同 IP地址的RP（相当于一个RP虚拟组，就像VRRP、HSRP中的虚拟组一样）。这些相同的IP地址都配置在loopback接口上，且这些RP之间在一个PIM-SM域内建立MSDP对等体关系，从而实现RP路径最优及负载分担。总体而言，通过Anycast RP，可以实现以下三方面的好处。

（1）RP路径最优：组播源向距离最近的RP进行注册，建立路径最优的SPT；接收者向距离最近的RP发起加入，建立路径最优的RPT。

（2）RP间的负载分担：每个RP上只需维护PIM-SM域内的部分源/组信息、转发部分的组播数据，从而实现了RP间的负载分担。

（3）RP间的冗余备份：当某RP失效后，原先在该RP上注册或加入的组播源或接收者会自动选择就近的RP进行注册或加入操作，从而实现了RP间的冗余备份。

如图12-29所示，在PIM-SM域内组播源Source1和Source2向组播组G发送组播数据，Receiver1和Receiver2是组播组G的成员。为了减轻单台RP的负荷，希望采用Anycast RP来实现多RP负载均衡。具体部署方法如下。

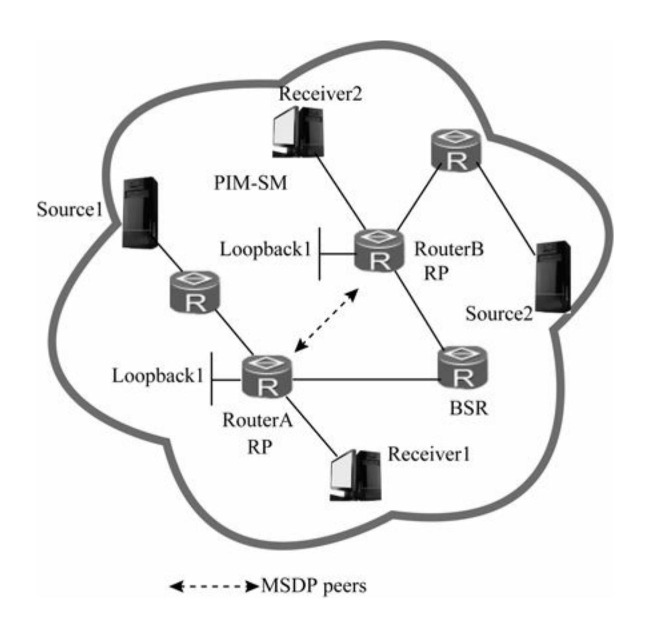


图12-29 Anycast RP典型应用示例

（1）在 PIM-SM 域内选取多台路由器（如图中的RouterA和RouterB）成为待选RP。

（2）在这些路由器上各准备一个Loopback接口，并配置相同的接口地址，如图中两路由器的

Loopback1接口。

（3）采用以下方式之一配置RP。

① 使用静态 RP：在全域的所有PIM-SM 路由器上配置静态 RP 的 IP 地 址，即上面在RouterA的Loopback1接口上配置的IP地址。

② 使用C-RP：在RouterA和RouterB上使用Loopback1接口配置C-RP，然后在网络中配置C-BSR，选举产生BSR，用于集中管理组播域中的RP、组播源注册和组成员加入消息。注意：C-RP的地址不能与C-BSR的地址相同。

（4）在RouterA和RouterB两个路由器之间配置建立PIM-SM域内的MSDP对等体 连接。但要注意，此时不能使用两路由器上Loopback1接口上的RP IP地址，而要使用路由器实际连接的物理接口IP地址。

12.4.4 组播源信息在域间的传递

各个PIM-SM域的RP之间配置了MSDP对等体关系后，MSDP对等体之间通过交互SA（Source Active）消息（SA消息中携带组播源DR在RP上注册时的（S，G）信息）在各MSDP对等体之间的信息传递，这样任意一个RP发出的SA消息能够被其他所有的RP收到。

如图12-30所示，PIM-SM网络被划分为4个PIM-SM域。PIM-SM1域内的组播源Source向组G发送数据。PIM-SM3域内的Receiver为组G成员，RP3和Receiver之间维护了一棵关于组G的RPT（RP-rooted Shared Tree，根RP共享树）。

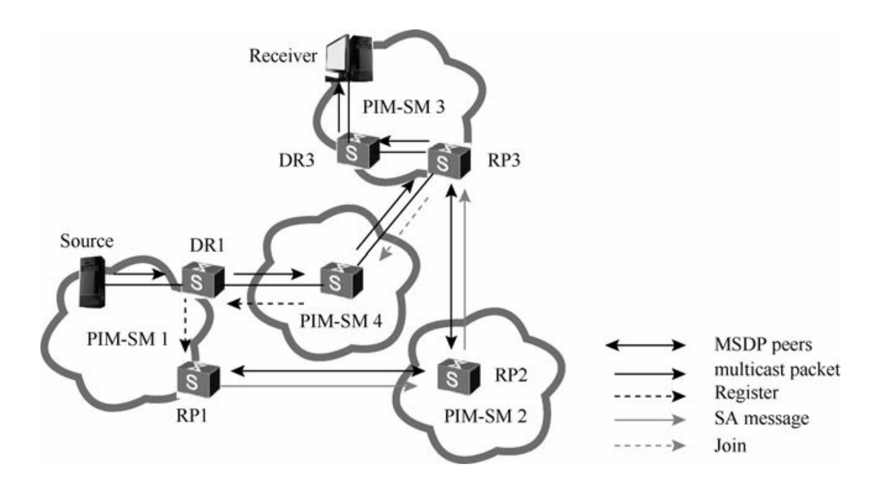


图12-30 MSDP实现域间组播示例

通过在RP1、RP2和RP3之间建立MSDP对等体关系，可以使Receiver（接收者）接收到Source发出的组播数据，具体过程如下。

（1）Source向组G发送组播数据。DR1将组播数据封装在Register消息中发给RP1。RP1作为源端RP，创建SA消息，携带Source的IP地址、组G地址和RP1地址发送给对等体RP2。

（2）RP2接收到该SA消息后，执行RPF（Reverse Path Forwarding，逆向路径转发）检查，通过后向RP3转发。

（3）RP3接收到该SA消息后，同样先执行RPF检查，由于RP3上存在（*，G）表项，表示本域内存在组G成员，RP3直接创建（S，G）表项。

（4）然后RP3向Source方向逐跳发送（S，G）加入消息，创建一条从Source到RP3的组播路径（源树）。组播数据沿源树到达RP3后，再沿RPT向接收者转发。

（5）接收者接收到组播数据后，自行决定是否发起SPT切换。

12.4.5 SA消息转发的控制

在MSDP协议中，SA（源激活）消息在MSDP对等体之间转发，除了RPF检查， 还可以配置各种转发策略的过滤，从而只接收和转发来自正确路径并通过过滤的 SA 消息，以避免SA消息传递环路。另外，可以在MSDP对等体之间配置MSDP全连接组（Mesh Group），以避免SA消息在MSDP对等体之间的泛滥。

1．SA消息的RPF检查规则

为了防止SA消息在MSDP对等体之间被循环转发，MSDP对接收到的SA消息执行RPF检查，在消息传递的入方向上进行严格的控制。不符合RPF规则的SA消息，将被丢弃。

RPF检查的主要规则：MSDP设备收到SA消息后，根据MRIB（Multicast RPF Routing Information Base，组播RPF路由信息库）确定到源RP（即创建该SA消息的RP）最佳路径的下一跳是哪个对等体。这个对等体也称为“RPF对等体”。如果发现SA消息是从RPF对等体发出的，则接收该SA消息，并向其他对等体转发。

此外，还有如下的一些RPF检查规则，SA消息在转发时必须遵守。

（1）发出SA消息的对等体就是源RP，则接收该SA消息并向其他对等体转发。

（2）接收从静态 RPF 对等体到来的 SA 消息。一台路由器可以同时与多个路由器建立 MSDP 对等体关系。用户可以从这些远端对等体中选取一个或多个配置为静态RPF对等体。

（3）如果一台路由器只拥有一个远端MSDP对等体，则该远端对等体自动成为RPF对等体，路由器接收从该远端对等体发来的SA消息。

（4）发出SA消息的对等体与本地路由器属于同一Mesh Group，则接收该SA消息。来自Mesh Group的SA消息不再向属于该Mesh Group的成员转发，但向该Mesh Group之外的所有对等体转发。

（5）到达源RP的路由需要跨越多个AS时，接收从下一跳AS（以AS为单位）中的对等体发出的SA消息，如果该AS中存在多个远端MSDP对等体，则接收从IP地址最高的对等体发来的SA消息。

2．MSDP全连接组（Mesh Group）

当网络中存在多个MSDP对等体时，很容易导致SA消息在对等体之间泛滥。同时， MSDP对等体对每一个到来的SA报文进行RPF检查，给系统造成很大的负担。将多个MSDP对等体加入同一个全连接组（Mesh Group），就可以大幅度减少在这些MSDP对等体之间传递的SA消息。

Mesh Group 成员可以都属于同一个PIM-SM域，也可以分布在多个PIM-SM域中；可以都位于同一个AS，也可以位于多个AS中。但属于同一个Mesh Group的所有成员之间必须两两建立 MSDP 对等体连接，并承认对方为该Mesh Group的成员。

如图 12-31 中的 RouterA、RouterB、RouterC 和 RouterD ， 加入同一个 Mesh Group，则必须在每台路由器上配置与其他三台路由器建立MSDP对等体关系。

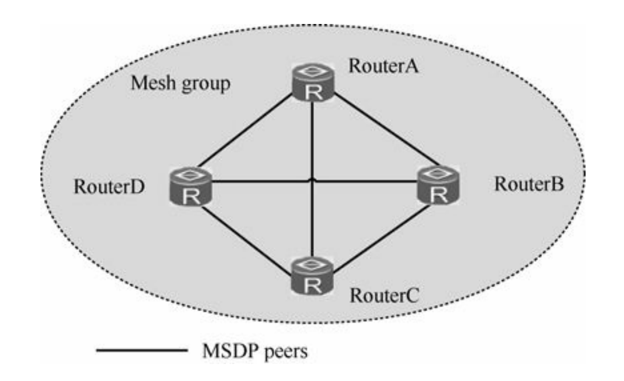


图12-31 Mesh Group内部成员之间的MSDP对等体连接示例

当Mesh Group内部成员接收到SA消息后，首先检查该SA消息的来源。

（1）如果该SA消息来自Mesh Group外部的某个MSDP对等体，则对该SA消息进行RPF检查。如果检查通过，向Mesh Group内其他所有成员转发。

（2）如果该SA消息来自Mesh Group内部成员，则不进行RPF检查，直接接收。同时也不再向Mesh Group内其他成员转发。

3．SA消息过滤

缺省情况下，MSDP不过滤SA消息，从一个域中发出的SA消息可以被传递到全网的 MSDP 对等体。然而有些 PIM-SM 域的（S，G）表项只适用于本域内转发，如一些本地组播应用使用了全局的组播组地址，或组播源用的是私网地址10.x.x.x。

如果不加过滤，这些（S，G）表项就会经过SA消息传递到其他MSDP对等体。针对这种情况，可以配置SA消息的过滤规则（一般使用ACL定义过滤的规则），并在创建、转发或接收SA消息时使用这些规则，就可以实现SA消息过滤。

12.4.6 MSDP的应用

通过前面的介绍，我们已经知道，MSDP既可以应用于不同PIM-SM域之间，又可应用于同一PIM-SM域之内；既可应用于不同AS的PIMS-SM域之间，又可应用于同一个AS的PIM域之间。下面分别简单介绍。

1．MSDP在AS内PIM-SM域间组播应用

组播源和组成员处于同一AS内不同PIM-SM域间，可以在RP上配置MSDP对等体，实现域间组播。如图12-32所示，在两个PIM-SM域的RP之间配置MSDP对等体，可以共享对方域内的组播源信息。

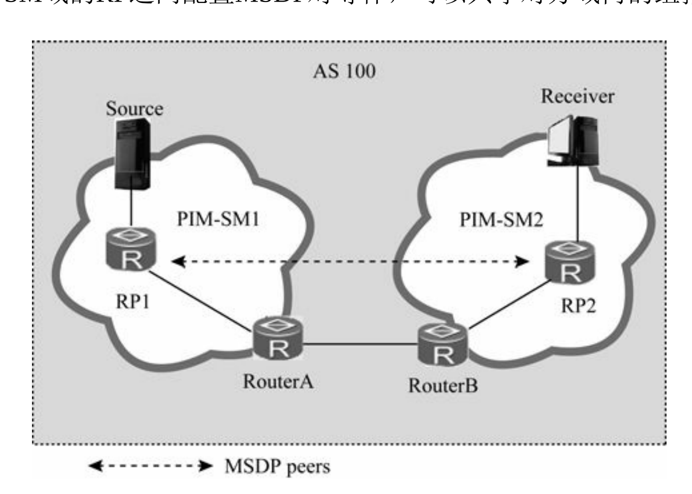


图12-32 MSDP在AS内PIM-SM域间的组播应用示例

2．MBGP在AS间PIM-SM域间组播的应用

MBGP也可应用于不同AS间的PIM-SM域间，此时在MSDP对等体之间建立MBGP对等体，可以保证SA消息顺利通过RPF检查。配置时将MBGP对等体和MSDP对等体建立在相同的接口上。

如图12-33所示，各个PIM-SM中的RP并不直接相连，且与ASBR不在同一台路由器上。在它们之间建立MSDP对等体后，为了保证SA消息顺利通过RPF检查，在这 些RP之间配置MBGP对等体。

（1）IBGP：在属于同一AS的RP之间建立IBGP对等体。例如：RP1和RP2、RP3和RP4。

（2）EBGP：在属于不同AS的RP之间建立EBGP对等体。例如：RP1和RP3、RP2和RP4。

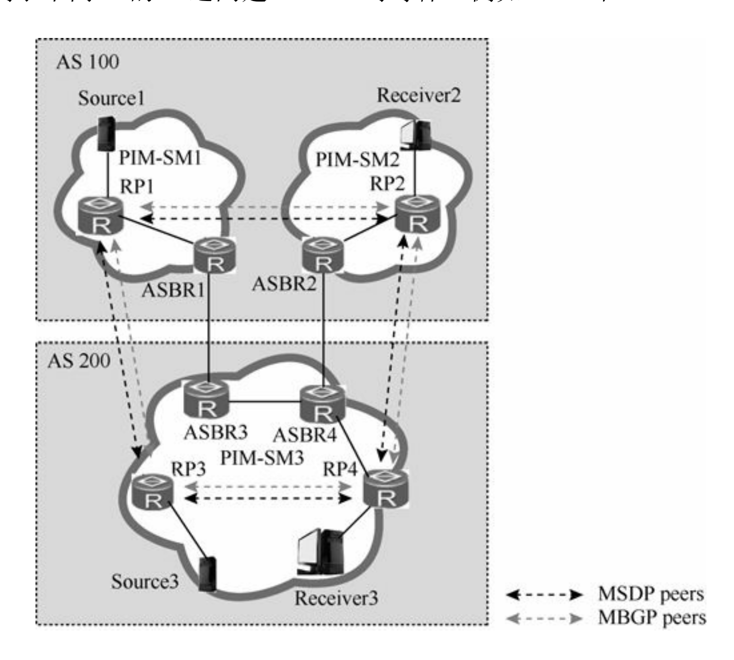


图12-33 MSDP在不同AS间的组播应用

3．使用静态RPF对等体实现AS间组播

在12.4.5节介绍的SA消息的RPF检查规则中提到，如果收到的SA消息是从“RPF对等体”发来的，则接收该消息并向其他对等体转发。如果在MSDP对等体之间手工指定对方互为静态RPF对等体，从静态RPF对等体收到的SA消息将不做RPF检查。

静态RPF配置不当容易引起SA消息环路，请慎重使用。通常在跨AS的MSDP对等体之间建立静态RPF对等体。如图12-34所示，在RP1和RP2之间配置MSDP对等体；在RP1和RP2上分别配置对方为自己的静态RPF对等体。

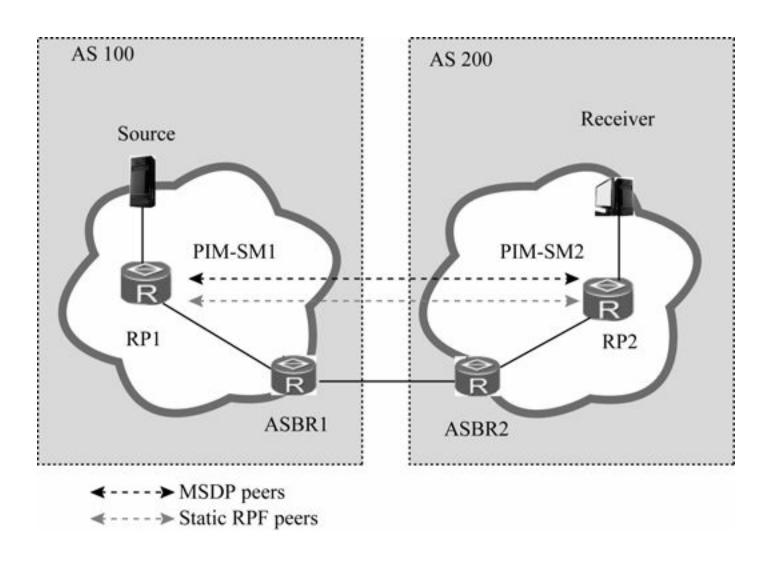


图12-34 使用静态RPF对等体实现AS间组播

4．Anycast RP

在12.4.3节中介绍到，MSDP也可以在PIM-SM域内通过Anycast RP功能得到应用，实现域内多个RP间的负载分担和容错。

如图12-35所示，Switch1和Switch2作为RP，两者之间建立MSDP对等体关系。借助MSDP对等体进行域内组播，接收者选择距离最近的RP发送加入消息以构建RPT树；组播源可以选择距离最近的RP进行注册，RP之间交互SA消息，共享源信息。RP加入以源端DR为根的SPT，引入组播数据；接收者接收到组播数据后自行决定是否发起SPT切换。

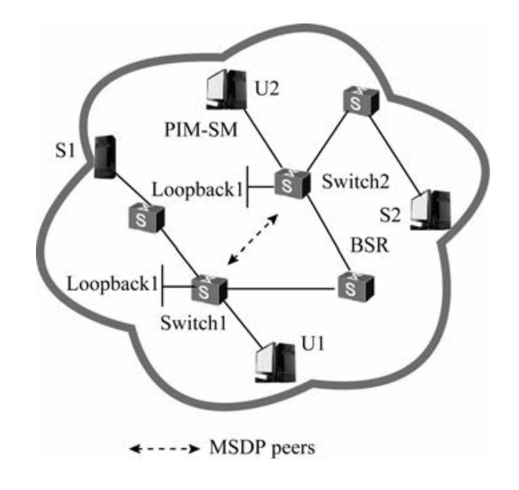


图12-35 MSDP在PIM-SM域内的Anycast RP应用

12.5 二层组播基础及工作原理

二层组播是指组播信息在数据链路层的按需分发。二层组播的基本原理是使二层设备可以识别组播组IP地址，建立组播组IP地址与端口对应关系的组播转发表，指导组播数据在数据链路层的转发。

12.5.1 二层组播概述

在很多情况下，组播报文要不可避免地经过一些二层交换设备，尤其是在局域网环境里，许多接入交换机都是二层的。如图12-36所示，在组播用户和三层组播设备Router之间，经过二层交换机Switch。

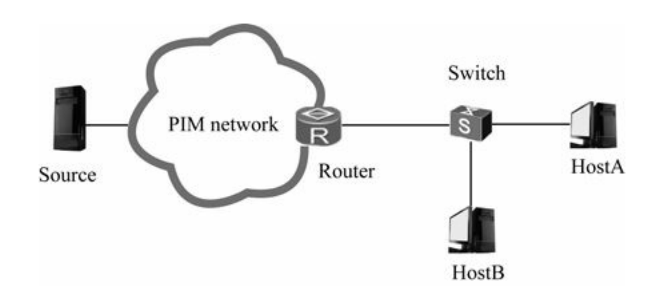


图12-36 二层组播网络示例

当Router将组播报文转发至Switch以后，Switch负责将组播报文转发给组播用户。由于组播报文的目的IP 地址为组播组 IP 地址，在二层设备上是学习不到这一类 MAC表项的，因此，组播数据报文就会在所有接口进行广播，和它在同一广播域内的组播成员和非组播成员都能收到组播数据报文。这样不但浪费了网络带宽，而且影响了网络信息安全。

通过以下这些二层组播技术，可以控制这种广播。

（1）IGMP Snooping/MLD Snooping：二层设备侦听组播用户和上游路由器之间的IGMP/MLD报文，建立二层组播转发表，控制组播数据报文转发。

（2） IGMP Snooping Proxy/MLD Snooping Proxy： IGMP Snooping/MLD Snooping 协议报文代理，可减少协议报文转发，降低上游设备的性能压力，节省上游网络带宽。

（3）组播VLAN：上游设备只把组播数据传送给一个指定VLAN，然后由此VLAN在二层设备上把组播流复制到其他VLAN中，可大大减少上游网络带宽的浪费。

（4）二层组播CAC（Call Admission Control，呼叫许可控制）：控制组播组的数量和带宽，避免出现接入带宽需求超出汇聚网络带宽的情况，保证大多数用户的服务质量。

（5）基于VLAN的可控组播：控制用户加入某个组播组的权限。当用户请求加入某个组播组时，二层设备必须对这个请求进行验证，拒绝非法或越权的请求。

此外，通过二层组播SSM Mapping，还可以使 IGMPv1和 IGMPv2版本的主机享受SSM的组播服务。

12.5.2 IGMP Snooping/MLD Snooping基本原理

IGMP Snooping/MLD Snooping是二层组播的基本功能，可以实现组播数据在数据链路层的转发和控制。当主机和上游三层设备之间传递的IGMP/MLD协议报文通过二层设备时，IGMP Snooping/MLD Snooping分析报文携带的信息，根据这些信息建立和维护二层组播转发表，从而指导组播数据在数据链路层按需转发。

IGMP Snooping用于 IPv4组播网络，MLD Snooping用于 IPv6组播网络。除了使用的地址和协议报文名称不同，二者实现原理相同。下文以 IGMP Snooping为例描述工作原理。

如图12-37所示，当组播数据从三层组播设备Router转发下来以后，处于接入边缘的二层设备Switch负责将组播信息转发给用户。

当二层设备没有运行 IGMP Snooping时，组播数据在二层被广播（因为二层交换机无法识别组播 IP 地址），包括非组播成员都会收到该组播报文；而当二层设备运行了IGMP Snooping后，已知组播组的组播数据不会在二层广播，而是会被组播发送给指定的接收者，没有加入对应组播组的用户不会收到组播报文。

因为使能 IGMP Snooping功能后，二层设备会侦听主机和上游三层设备之间交互的IGMP报文，通过分析报文中携带的信息（报文类型、组播组地址、接收报文的接口等），建立和维护二层组播转发表，从而指导组播数据在数据链路层按需转发。

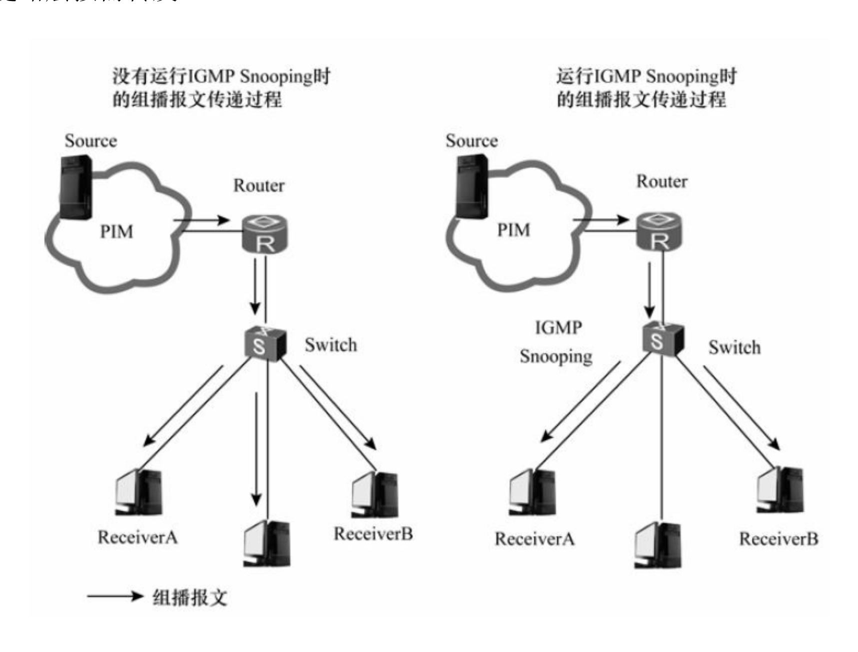


图12-37 二层设备运行 IGMP Snooping前后组播报文传递对比

1．端口角色

在二层组播 IGMP Snooping协议中涉及几种端口角色，下面以图12-38为例进行介绍。

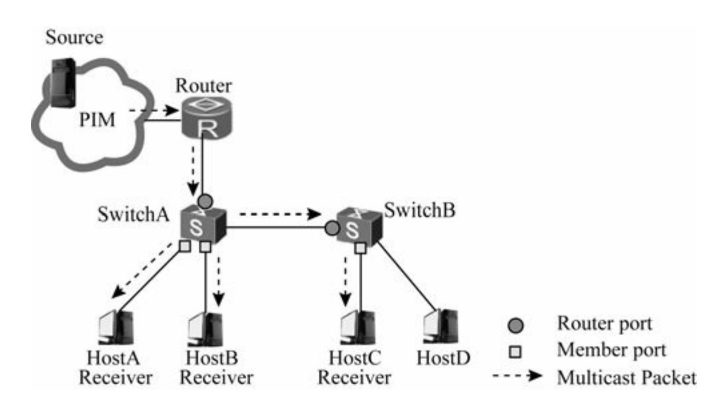


图12-38 IGMP Snooping相关端口角色

IGMP Snooping中相关端口角色如表 12-8所示。

表12-8 IGMP Snooping中的端口角色

端口角色	说明
路由器端口（Router Port）：指二层设备上连接组播路由器的接口，而不是指路由器上的接口。如图中 SwitchA 和 SwitchB 上用圆圈表示的接口	二层设备从此接口接收组播数据报文。由协议生成的路由器端口叫做动态路由器端口。收到源地址不为 0.0.0.0 的 IGMP 普遍组查询报文或 PIM Hello 报文（三层组播设备 PIM 接口向外发送的用于发现并维持邻居关系的报文）的接口都将被视为动态路由器端口。手工配置的路由器端口叫做静态路由端口。
成员端口（Member Port），如 SwitchA 和 SwitchB 上用方框表示的接口	又称组播组成员端口，二层设备往此接口发送组播数据报文。 由协议生成的成员端口叫做动态成员端口。收到 IGMP Report 报文的接口，二层设备会将其标识为动态成员端口。由手工配置的成员端口叫做静态成员端口

二层交换机上的路由器端口和成员端口是二层组播转发表项中的一个重要信息，代表出接口。其中路由器端口相当于上游接口，成员端口相当于下游接口。通过协议报文学习到的端口，对应的为动态表项；而手工配置的端口，对应的为静态表项。

除了出接口外，每条二层组播转发表项还包括组播组地址和VLAN编号。组播组地址，可以为组播IP地址，也可以为组播IP地址映射后的组播MAC地址。按照IP地址转发的模式可以避免MAC地址转发模式中的地址重复问题。

VLAN编号指定了二层广播域范围。如果使用了组播VLAN功能，入VLAN编号为组播VLAN的编号，出VLAN编号为主机所在的用户VLAN编号。否则入VLAN编号和出VLAN编号均为主机所在VLAN的编号。有关组播VLAN将在本章后面详细介绍。

2．工作机制

在二层设备运行了 IGMP Snooping后，收到不同的 IGMP协议报文会进行不同的处理，并在此过程中建立起二层组播转发表项，具体如表12-9所示。

表12-9 IGMP Snooping对不同报文的处理方式

IGMP 工作阶段	二层设备收到的报文类型及处理方式
普遍组查询：IGMP 查询器定期向本地网段内的所有主机与路由器（目的 IP 地址为 224.0.0.1）发送 IGMP 普遍组查询报文，以查询该网段有哪些组播组的成员	此时收到的是 IGMP 普遍组查询报文。二层设备会向 VLAN 内除接收接口外的其他所有接口转发，并对接收接口做如下处理。 • 如果路由器端口列表中已包含该动态路由器端口，则重置老化定时器。收到 IGMP 普遍组查询报文时，动态路由器端口的老化定时器缺省为 180s，可以通过命令行配置 • 如果路由器端口列表中尚未包含该接口，则将其添加进去，并启动老化定时器
成员关系报告：成员收到 IGMP 普遍组查询报文后，回应 IGMP Report 报文；成员主动向 IGMP 查询器发送 IGMP Report 报文以声明加入该组播组	此时收到的是 IGMP Report 报文。二层设备会向 VLAN 内所有路由器端口转发。从报文中解析出主机要加入的组播组地址，并对接收接口做如下处理。 • 如果不存在该组对应的转发表项，则创建转发表项，将该接口作为动态成员端口添加到出接口列表中，并启动老化定时器 • 如果已存在该组对应的转发表项，但出接口列表中未包含该接口，则将该接口作为动态成员端口添加到出接口列表，并启动老化定时器 • 如果已存在该组所对应的转发表项，且出接口列表中已包含该动态成员端口，则重置其老化定时器
成员离开组播组：运行 IGMPv2 或 IGMPv3 的成员发送 IGMP Leave报文，以通知组播路由器自己离开了某个组播组；IGMP 查询器收到 IGMP Leave 报文后，从中解析出组播组地址，并通过接收接口向该组播组发送 IGMP 特定组查询报文/IGMP 特定源组查询报文	如果收到的是 IGMP Leave 报文，则二层设备会判断离开的组是否存在对应的转发表项，以及转发表项出接口列表是否包含报文的接收接口。 • 如果不存在该组对应的转发表项，或者该组对应转发表项的出接口列表中不包含接收接口，二层设备不转发该报文，将其直接丢弃 • 如果存在该组对应的转发表项，且转发表项的出接口列表中包含该接口，二层设备会将报文向 VLAN 内所有路由器端口转发 对于 IGMP Leave 报文的接收接口（假定为动态成员端口），二层设备在其老化时间内按如下规则处理。 • 如果从该接口收到了主机响应该特定组查询的 IGMP Report 报文，表示接口下还有该组的成员，于是重置其老化定时器 • 如果没有从该接口收到主机响应特定组查询的 IGMP Report 报文，则表示接口下已没有该组成员，则在老化时间超时后，将接口从该组的转发表项出接口列表中删除
	如果收到的是 IGMP 特定组查询报文/IGMP 特定源组查询报文，则二层设备会向 VLAN 内除接收接口外的其他所有接口转发

此外，当二层设备收到PIM Hello报文时，会向VLAN内除接收接口外的其他所有接口转发，并对接收接口做如下处理。

（1）如果路由器端口列表中已包含该动态路由器端口，则重置老化定时器。

（2）如果路由器端口列表中尚未包含该接口，则将其添加进去，并启动老化定时器。

收到PIM Hello报文时，动态路由器端口的老化时间为Hello报文中Holdtime字段的值。

如果是静态配置路由器端口，二层设备收到 IGMP Report和Leave报文也会向静态路由器端口转发。如果配置了静态成员端口，则转发表项中会添加该接口为出接口。

当二层设备上建立了二层组播转发表以后，在接收到组播数据报文时会依据报文所属VLAN和报文的目的地址（即组播组地址）查找转发表项是否存在对应的“出接口信息”。如果存在，则将报文发送到所有组播组成员端口；如果不存在，则丢弃该报文或将报文在VLAN内广播。

12.5.3 IGMP Snooping Proxy/MLD Snooping Proxy基本原理

为了减少上游三层设备收到的 IGMP Report/MLD Report报文和 IGMP Leave/MLD Done报文的数量，可以在二层设备上部署 IGMP Snooping Proxy/MLD Snooping Proxy功能，使其能够代理下游主机来向上游设备发送成员关系报告报文。配置了 IGMP Snooping Proxy/MLD Snooping Proxy功能的设备称为 IGMP Snooping/MLD Snooping代理，在其上游设备看来，它就相当于一台主机；在其下游设备看来，它相当于一台查询器。

IGMP Snooping Proxy用于 IPv4组播网络，MLD Snooping Proxy用于 IPv6组播网络。除了使用的地址和协议报文名称不同，二者实现原理相同。下文仅以 IGMP Snooping Proxy为例描述工作原理。

1．在组播报文传递上的改变

如图12-39所示，当Switch上运行 IGMP Snooping时，Switch对上游Router的Query报文和下游主机的Report/Leave报文都是原封不动地转发。当网络中存在大量用户主机时，冗余的IGMP报文给上游设备带来处理压力。

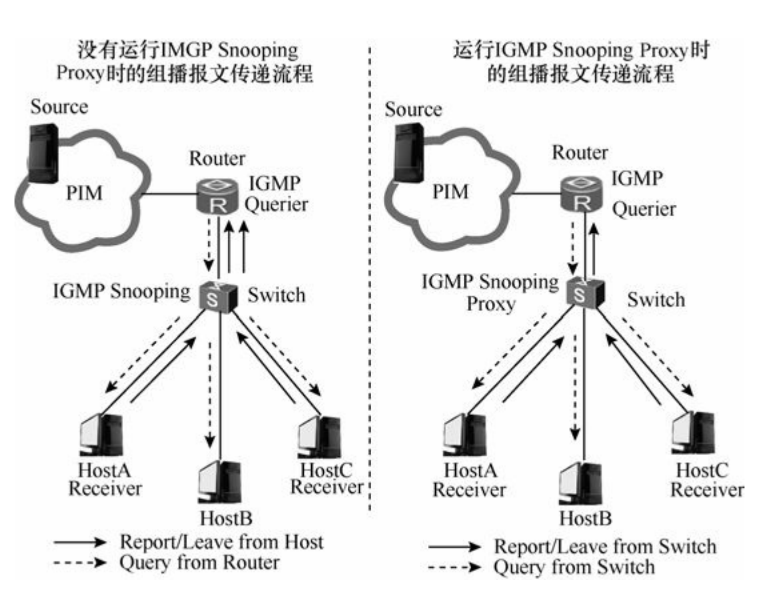


图12-39 IGMP Snooping Proxy运行前后的组播报文传递流程

当Switch上配置 IGMP Snooping Proxy时，Switch可以终结上游的 IGMP Query报文，并且自己构造Query报文向下游主机发送；终结下游主机的 IGMP Report/Leave报文，并自己构造统一的Report/Leave报文向上游发送。

部署 IGMP Snooping Proxy后，三层设备会感知到下面只有一个用户，二层设备直接跟下游用户和三层设备进行对话，而不再是一个完全透明的转发角色。IGMP Snooping Proxy可有效减少 IGMP协议在网络中的交互程度，节约带宽；有效屏蔽来自下游主机的大量协议报文，并接管了对主机的查询器功能，分担上游三层设备的性能负荷。

2．工作机制

运行 IGMP Snooping Proxy的设备会参与二层组播转发表的建立和维护，依据转发表向有需要的用户主机发送组播数据。代理设备对各种IGMP报文的处理方式如表12-10所示。

表12-10 IGMP Snooping Proxy对接收到的 IGMP报文的处理方式

IGMP 报文类型	处理方式
IGMP 普遍组查询报文	向本 VLAN 内除接收接口以外的所有接口转发；同时根据本地维护的组成员关系生成 Report 报文，向所有路由器端口发送
IGMP 特定组查询报文/IGMP 特定源组查询报文	若该组对应的转发表项中还有成员端口，则向所有路由器端口回复该组的 Report 报文
IGMP Report 报文	（1）如果不存在该组对应的转发表项，则创建转发表项，将接收接口作为动态成员端口添加到出接口列表中，并启动其老化定时器，然后向所有路由器端口发送该组的 Report 报文 （2）如果已存在该组对应的转发表项，且其出接口列表中已包含该动态成员端口，则重置其老化定时器 （3）如果已存在该组对应的转发表项，但其出接口列表中不包含该接收接口，则将该接口作为动态成员端口添加到出接口列表中，并启动其老化定时器
IGMP Leave 报文	向接收接口发送针对该组的特定组查询报文。只有当删除某组播组对应转发表项中的最后一个成员端口时，才会向所有路由器端口发送该组的 Leave 报文

12.5.4 二层组播SSM Mapping

通过前面的学习已知，SSM（指定源组播）与 ASM（任意源组播）技术相比就是SSM 可以指定组播源，具有更好的安全性。但从本章前面 12.2 节的介绍可知，在三层IGMP和MLD组播协议中，只有IGMPv3或MLDv2版本协议支持SSM，如果用户主机只能运行IGMPv1/IGMPv2或MLDv1，则不能直接使用SSM模型。这时，为了使其能够使用SSM服务，则需要借助SSM Mapping功能。

同样，在二层组播中也可通过SSM Mapping实现对SSM服务的支持，但是目前只实现了 IPv4组播网络中的二层SSM Mapping，即基于 IGMP Snooping的SSM Mapping。 通过在 IGMP Snooping协议的二层设备上静态配置SSM地址的映射规则，可将 IGMPv1和 IGMPv2 Report报文中的（*，G）信息转化为对应的（S，G）信息，以提供SSM组播服务。S 表示组播源，G 表示组播组，*表示任意组播源。缺省情况下，SSM 组播组IP地址范围为232.0.0.0～232.255.255.255。

如图12-40所示，3个接收者（Recevier）运行不同的IGMP版本，且HostB和HostC无法升级到IGMPv3。如果要为该网段中的所有主机提供SSM服务，则必须在二层设备Switch上使用二层组播SSM Mapping功能。

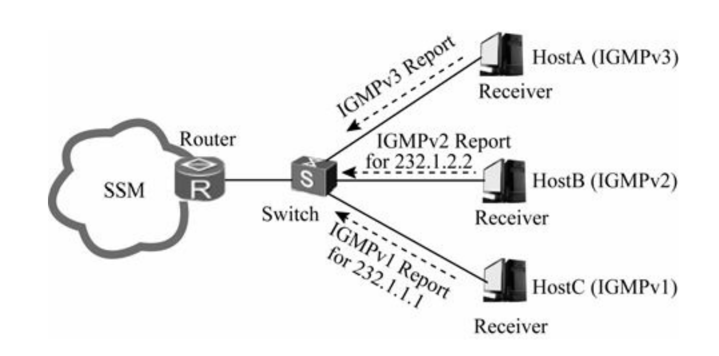


图12-40 二层组播SSM Mapping应用示例

这时，如果在Switch上配置了如下组播组地址和组播源地址映射关系。

（1）232.1.1.0/24→10.10.1.1

（2）232.1.2.0/24→10.10.2.2

（3）232.1.3.0/24→10.10.3.3

经过映射后，Switch在收到HostB和HostC的成员报告报文时，首先判断报文携带的组地址是否在SSM范围内，经过映射后则肯定在SSM范围内，此时就可以根据配置的映射规则生成如下所示的组播表项。

（1）232.1.1.1（来自HostC）：（10.10.1.1，232.1.1.1）

（2）232.1.2.2（来自HostB）：（10.10.2.2，232.1.2.2）

如果 Report 报文携带的组地址在 SSM 范围内，但是 Switch 上没有对应的 SSM Mapping规则，则无法提供SSM服务，丢弃该报文。如果Report报文携带的组地址不在SSM范围内，则只提供ASM服务。

12.5.5 组播VLAN

在 12.5.1节介绍的 IGMP Snooping/MLD Snooping二层组播侦听功能很好地弥补了组播数据在二层广播网络只能进行广播的不足。但是这种功能仍是基于一个广播域，即基于 VLAN 来实现的。如果不同 VLAN 的用户有相同的组播数据需求时，上游路由器仍然需要发送多份相同报文到不同 VLAN 中。这时就得使用组播 VLAN 功能了。它实现了组播路由器只需发送一份数据就可在二层网络设备上进行跨VLAN组播复制，大大减轻了组播路由器和网络的负担。

1．组播VLAN基本原理

如图 12-41 所示，属于不同 VLAN（VLAN2 和 VLAN3）的用户需要接收相同的 组播流，上游路由器RouterA就会把组播数据在每个VLAN内都复制一份然后发送给下游交换机SwitchA。这样，既造成了路由器与二层设备之间带宽的浪费，也增加了路由器额外的负担。

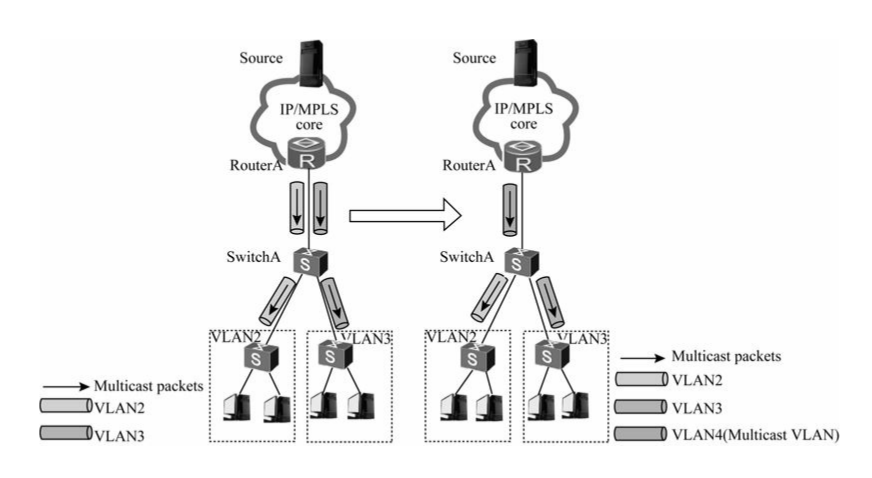


图12-41 配置组播VLAN功能前后的组播报文传递对比

通过在二层设备上配置组播 VLAN 功能就可以解决这个问题。如在图中的 Switch上部署了组播VLAN功能后，上游路由器不必在每个用户VLAN（VLAN2和VLAN3）内都复制一份组播流，而组播路由器只需向组播VLAN（VLAN4）内复制一份数据发送给二层设备。这样就避免了组播流在上游设备的重复复制，不仅节省了网络带宽，而且减轻了上游路由器的负担。

在这里涉及两种VLAN：组播VLAN和用户VLAN。"组播VLAN"是网络侧接口（也就是连接组播路由器的接口）所属VLAN，用于实现组播流的汇聚；"用户VLAN"是用户侧接口（也就是连接组播接收者的接口）所属VLAN，用于接收组播VLAN复制、发送的组播数据副本。一个组播 VLAN 可以绑定多个用户VLAN，但一个用户 **VLAN**只能加入一个组播**VLAN** 。

2．组播VLAN的扩展

通常情况下，组播VLAN的应用场景是通过组播VLAN将相同的组播数据复制分发到不同的用户 VLAN中来节省网络带宽。但是，有时候也会存在单个用户VLAN需要接收多个组播VLAN的组播数据的场景。为此，组播VLAN进行了如下的扩展。

（1）组播VLAN多对多

"组播VLAN多对多"是对传统的"组播VLAN一对多"（即多个用户VLAN可以加入一个组播 VLAN，但是一个用户 VLAN 不能加入多个组播 VLAN）的补充。如图12-42所示，用户VLAN（UVLAN）需要接收来自两个组播VLAN（MVLAN1、MVLAN2）的组播数据。通过组播 VLAN 多对多功能实现一个用户同时接收两个组播 VLAN 数据的过程如下。

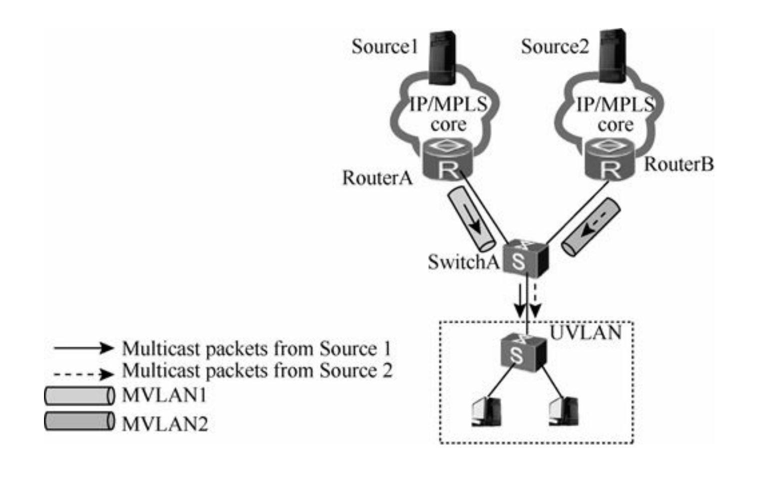


图12-42 组播VLAN多对多示例

① 在用户VLAN上使能静态组播流触发功能。

② 在组播VLAN上分别配置Source1、Source2的静态组播流。

③ 用户VLAN分别加入两个组播VLAN。

从上述过程可以看出，组播 VLAN 多对多主要是依靠静态组播流触发机制+组播VLAN一对多功能共同来完成。

（2）基于接口的组播VLAN

在组播VLAN中还会出现这样一种场景：网络中的组播业务批发给了多个ISP，单个用户VLAN中不同用户定制了不同的ISP所提供的组播业务。这时候如果再使用组播VLAN多对多功能，就会造成只定制某ISP提供的组播业务的用户能够接收到其他ISP提供的组播业务。

为了保证相同用户VLAN中的不同用户之间的组播业务隔离，可通过为不同的ISP分配一个VLAN作为组播VLAN，然后基于接口将组播VLAN与用户VLAN进行绑定来实现，这就是基于接口的组播VLAN。这样，ISP对应的组播VLAN就与{用户接入接口、用户VLAN}形成了一一映射的关系，用户接入接口就不会向下游转发未绑定的组播VLAN的数据。但只有 **IGMP Snooping**支持基于接口的组播**VLAN**。

如图12-43所示，组播业务批发给了ISP1、ISP2两个服务商，用户VLAN中的Host1和 Host2 定制的是ISP1 提供的组播服务，而Host3和Host4定制的是ISP2提供的组播服务。为了使两个ISP提供的组播数据不会发送到所有的用户主机上，向ISP1、ISP2 分别分配一个组播 VLAN （ MVLAN1、MVLAN2），将Hos1和Host2的接入接口与MVLAN1绑定，将Host3和Host4的接入接口与MVLAN2绑定。这样，ISP1提供的组播数据只向Host1和Host2发送，ISP2提供的组播数据只向Host3和Host4发送。

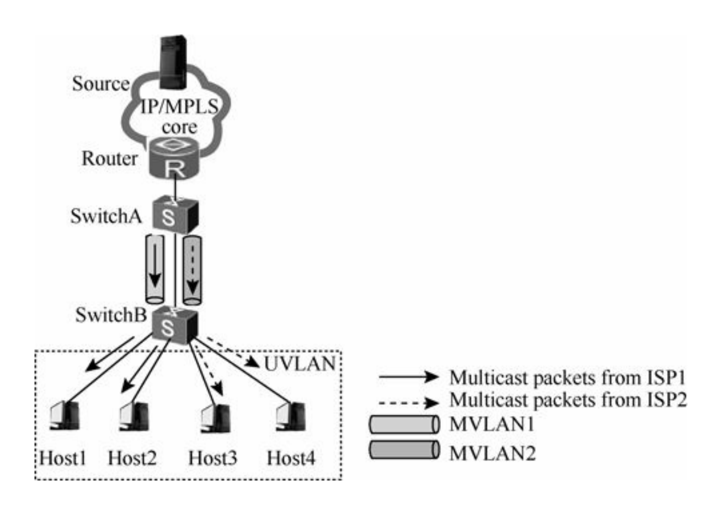


图12-43 基于接口的组播VLAN示例

12.6 组播路由管理

本节介绍的“组播路由管理”主要包括组播路由和转发、RPF检查、组播静态路由和组播负载分担这几个方面。通过组播路由可以控制组播报文的转发，通过RPF可以组播转发路径的检测和维护，通过单播路由、MBGP路由、组播静态路由可实现RPF检查，通过组播负载分担可实现不同转发路径下的流量分担。

12.6.1 组播路由和转发

“组播路由和转发”与“单播路由和转发”类似，首先，每个组播路由协议（典型代表为PIM和MBGP）都各自建立并维护了一张协议路由表（称之为“组播协议路由表”，这是组播路由的基础路由信息）。这就相当于单播路由中的各单播路由协议路由表，如RIP路由表、OSPF路由表、BGP路由表等。其次，各“组播协议路由表”的组播路由信息经过综合又形成一个总的“组播路由表”（Multicast Routing-Table），就相当于各单播路由协议的路由表最终形成了统一的IP路由表。最后，组播路由器再根据“组播路由和转发”策略，从组播路由表中选出最优的组播路由下发到“组播转发表”（Multicast Forwarding-Table，相当于单播路由中的三层转发表）中，直接用于控制组播数据的转发。通过“组播转发表”整个网络建立了一条以组播源为根，组成员为叶子的一点到多点的组播转发路径。

从以上的分析可以看出，在整个组播路由中至少包括以下三张“表”：组播协议路由表、组播路由表、组播转发表。另外在二层组播中，运行IGMP/MLD协议的路由器上还要维护两张“表”，即“IGMP/MLD组项”和“IGMP/MLD路由项”。下面以IPv4网络为例，介绍各“表”信息在实现组播路由和转发中所起的作用。

1．IGMP组和路由表

“IGMP组和路由表”包括“IGMP组表”和“IGMP路由表”两部分。IGMP组表是由用户主机发送的IGMP加入报文触发创建的，用于维护组成员加入信息并通知组播路由协议创建相应的（*，G）表项。只要设备接口使能了IGMP协议，并收到组成员加入报告报文就会为每个接口维护一个组加入信息表项。IGMP组表项可以通过display igmp group命令查看，具体形式如下所示。IGMP组表项中主要字段说明如表12-11所示。

```
<HUAWEI>display igmp group
Interface group report information of VPN-Instance: public net
  GigabitEthernet1/0/0 (10.1.6.2):
    Total 1 IGMP Group reported
  Group Address    Last Reporter    Uptime      Expires
  225.1.1.2        10.1.6.10        00:02:04    00:01:17
```

表12-11 IGMP组表项主要字段说明

字段	说明
Group Address	显示对应运行了 IGMP 协议的路由器接口所加入的组播组 IP 地址
Last Reporter	显示最后一次从该接口上发送组成员加入报告报文的用户 IP 地址
Uptime	显示最近一次从该接口收到组播报文至今的时间
Expires	显示对应组表项将被老化的时间，即该组表项在多长时间后将从 IGMP 组表中清除

从上表可以看出，IGMP组表项是由“组播组IP地址”、“最近一次从接口上发送组 成员加入报文的组成员IP地址”、“最近一次更新至今的时间”和“表项即将过期的时间”4 大部分组成。主要是前面两部分，通过这个表项的信息，就可以通过上层的组播路由协议为对应的组成员创建对应的组表项。

“IGMP路由表项”也是由IGMP协议维护的，但它只有在接口没有使能**PIM**协议才会存在 。它的作用主要是用来扩展组播路由表项的出接口。IGMP 路由表项可通过display igmp routing-table命令查看，具体形式如下。IGMP路由表项中主要字段说明如表12-12所示。

```
<HUAWEI >display igmp routing-table
Routing table of VPN-Instance: public net
  Total 1 entry
00001. (*, 225.1.1.1)
      List of 1 downstream interface
      GigabitEthernet1/0/0 (20.20.20.1),
```

Protocol: IGMP

表12-12 IGMP路由表项主要字段说明

字段	说明
00001. (*，225.1.1.1)	表示 IGMP 路由表中的第 00001 号表项，(*，225.1.1.1)，表示为 ASM 模式的路由表项，仅指定是组播组 IP 地址
List of 1 downstream interface	表示以上 IGMP 路由表项中的下游接口列表
Protocol: IGMP	表示生成以上下游接口的协议类型，此处显示为 IGMP，表明当前接口没有使能 PIM 协议。如果使能了 PIM 协议，则以 PIM 协议优先，显示的是 PIM

从上表可以看出，IGMP 路由表项主要用途就是指定对应组播组的下游组成员接口列表，以便组播报文可以从这些出接口上发送给组播成员。

2．组播协议路由表

“组播协议路由表”是组播路由器在运行各种组播路由协议（典型代表就是 PIM）时由各个协议自己维护的表项，是组播路由和转发的基础，也是对应组播路由协议自己进行组播数据路由的依据。就像单播路由中的RIP、OSPF等路由协议在同路由协议的网络中主要还是采用各自的路由表一样。PIM路由表项可通过display pim routing-table命令查看，具体形式如下。PIM路由表项中主要字段说明如表12-13所示。

```
<HUAWEI>display pim routing-table
VPN-Instance: public net
Total 0 (*, G) entry; 1 (S, G) entry
(172.168.0.12, 227.0.0.1)
    RP: 2.2.2.2
    Protocol: pim-sm, Flag: SPT LOC ACT
    UpTime: 02:54:43
    Upstream interface: GigabitEthernet1/0/0
        Upstream neighbor: NULL
        RPF prime neighbor: NULL
    Downstream interface(s) information:
    Total number of downstreams: 1
          1: GigabitEthernet2/0/0
             Protocol: pim-sm, UpTime: 02:54:43, Expires: 00:02:47
```

表12-13 PIM路由表项主要字段说明

字段	说明
（172.168.0.12，227.0.0.1）	以（S，G）格式显示组播协议路由表项
RP: 2.2.2.2	显示对应组播协议路由表项的 RP 地址，只有协议类型为 PIM-SM 时才会有此显示信息
Protocol: pim-sm	显示对应组播协议路由表项运行的组播协议类型和工作模型
UpTime: 02:54:43	显示对应组播协议路由表项已存在的时间
Flag: SPT LOC ACT	显示对应组播路由协议表项的标志
Upstream interface: GigabitEthernet1/0/0	显示对应组播路由协议表项的上游接口
Upstream neighbor: NULL	显示对应组播路由协议表项的上游邻居。NULL 表示不存在上游邻居
RPF prime neighbor: NULL	显示对应组播路由协议表项的 RPF 邻居。NULL 表示不存在 RPF 邻居
Downstream interface(s) information:	显示对应组播路由协议表项的下游接口信息
Expires: 00:02:47	显示对应组播路由协议表项的下游接口老化时间

从上表可以看出，组播路由协议表项中包括的信息比较丰富，其中主要包括组播组（包括组播源IP地址和组播组IP地址）、RP地址、上游接口、上游邻居、RPF邻居和下游接口。这些信息就构成了对应组播路由协议自己的组播转发路径基本信息。同时它也可为下面即将介绍的组播路由表提供基本信息。

3．组播路由表

“组播路由表”是组播路由管理模块生成的路由表，对应于单播路由中的IP路由表，是通过综合各种组播路由协议表信息而形成的。如果组播路由管理支持多种组播协议，那这里应该能看到多种组播路由协议生成的优选路由信息。但目前PIM DM和PIM SM不能同时运行。组播路由表的主要功能就是创建组播转发表。组播路由表项信息可通过display multicast routing-table命令查看，具体形式如下。组播路由表项中主要字段说明如表12-14所示。

```
<HUAWEI>display multicast routing-table
Multicast routing table of VPN-Instance: public net
Total 1 entry
00001. (172.168.0.2, 227.0.0.1)
      Uptime: 00:00:28
      Upstream Interface: GigabitEthernet1/0/0
      List of 2 downstream interfaces
          1: GigabitEthernet2/0/0
          2: GigabitEthernet3/0/0
```

表12-14 组播路由表项主要字段说明

字段	说明
00001. (172.168.0.2，227.0.0.1)	表示第 00001 号表项，此表项是（S，G）形式，是属于指定源的组播路由表项
UpTime: 02:54:43	显示组播路由表项最近一次更新至今的时间
Upstream Interface: GigabitEthernet1/0/0	显示对应组播路由表项的上游接口

从上表可以看出，“组播路由表”信息比较简单，主要包括组播组和上游接口。它主要用于生成下面即将介绍的“组播转发表”。

4．组播转发表

“组播转发表”是路由管理模块依据“组播路由表”信息生成的用于指导组播数据实际转发的表项，通常称为MFIB（组播转发信息库），这张表项与单播中FIB表的功能是一样的，用于指导组播数据转发。组播转发表通过display multicast forwarding-table命令可以查看，具体形式如下。组播转发表项中主要字段说明如表12-15所示。

```
<HUAWEI>display multicast forwarding-table
Multicast Forwarding Table of VPN-Instance: public net
Total 1 entry, 1 matched
00001. (172.168.0.2, 227.0.0.1)
      MID: 0, Flags: 0x0:0
      Uptime: 00:08:32, Timeout in: 00:03:26
      Incoming interface: GigabitEthernet1/0/0
      List of 1 outgoing interfaces:
        1:    GigabitEthernet2/0/0
```

Activetime: 00:23:15

Matched 38264 packets(1071392 bytes), Wrong If 0 packets

Forwarded 38264 packets(1071392 bytes)

表12-15 组播转发表项主要字段说明

字段	说明
00001. (172.168.0.2，227.0.0.1)	表示第 00001 号表项，此表项是（S，G）形式
Flags: 0x0:0	表示对应组播转发表项的标志
MID: 0	表示对应组播转发表项在 MFIB 表中的唯一标识，用于快速检索组播转发表
UpTime: 02:54:43	显示对应组播转发表项最近一次更新至今的时间
Timeout in: 00:03:26	显示对应组播转发表项的超时时间，即还有多长时间该表项将从转发表中清除
Incoming interface: GigabitEthernet1/0/0	显示对应组播转发表项的入接口
List of 1 outgoing interfaces:	显示对应组播转发表项的出接口列表
Activetime: 00:23:15	显示对应出接口已存在时间
Matched 38264 packets(1071392 bytes)	显示匹配对应组播转发表项的报文数目
Wrong If 0 packets	显示对应组播转发表项中从错误接口进入的报文数目
Forwarded 38264 packets(1071392 bytes)	显示对应组播转发表项中已转发的报文数目

从上表可以看出，“组播转发表”中包括的信息主要用于控制组播报文的转发，包括了各组播路由协议自己的路由表中的一些基本信息（如组播组、入接口列表、出接口列表）和报文转发统计信息。组播转发表主要用于RPF检查（因为它包括了RPF检查过程中必须要检查的“入接口”信息），组播报文的路由最终仍是由对应的组播协议路由表来完成的。

12.6.2 RPF检查

在单播路由与转发中，单播报文是沿着一条点到点的路径传输的，路由器只需要考虑报文需要到达的位置（即目的地址），就知道应该从哪个接口转发出去。组播路由与转发则不同，因为组播报文的目的地址为组播地址，只是标识了一组接收者，并不是一个可以唯一确定其位置的IP地址，所以无法通过“目的地址”来找到接收者的位置。但是组播报文的来源位置（即源地址）是可确定的，所以组播报文的转发路径可根据其源地 址的逆向查找来确保转发路径的正确性。具体过程如下。

（1）组播路由器在收到一份组播报文后，会根据报文的源地址通过单播路由表查找到达“报文源”的路由，查看到“报文源”的路由表项的出接口是否与收到该组播报文的入接口一致。

（2）如果一致，则认为该组播报文是从正确的接口到达，从而保证了整个转发路径的正确性和唯一性。

（3）如果不一致，则认为该组播报文非法，将被直接丢弃。

以上这个过程也就是我们本节要介绍的“RPF”（逆向路径转发）检查。检查通过后的“正确的接口”通常被称为RPF接口。下面再具体介绍RPF检查过程。

1．RPF检查过程

在RPF检查的过程中除依据单播路由表外，还会依据MBGP路由表、组播静态路由表。且缺省情况下，组播静态路由、MBGP 路由和单播路由的优先级是依次降低的。当路由器收到一份在这全个路由表中均存在对应路由表项的组播报文后，按照以下原则进行具体检查。

（1）首先，通过报文源地址（是单播IP地址）分别从单播路由表、MBGP路由表和组播静态路由表中各选出一条最优路由。单播路由、MBGP路由的出接口为RPF接口，下一跳为RPF邻居；而组播静态路由实际上属于手工配置的组播路由，已经明确指定了RPF接口与RPF邻居。

（2）根据以下原则从这三条最优路由中选择一条作为RPF路由。

① 如果配置了按照最长匹配原则（也是掩码最长，路由最精确）来选择路由，则从这三条路由中选出最长匹配的那条路由，如果这三条路由的掩码长度一样，则选择优先级最高的那条路由；如果它们的优先级也相同，则按照组播静态路由、MBGP路由、单播路由的先后顺序进行选择。

② 如果没有配置按照最长匹配原则来选择路由，则从这三条路由中选出优先级最高的那条路由；如果它们的优先级相同，则也按照组播静态路由、MBGP路由、单播路由的先后顺序进行选择。

（3）最后，路由器会将报文的入接口与上面最终选择的 RPF 路由的 RPF 接口进行比较。如果一致则RPF检查通过，表明该报文来源路径正确，会将其向下游转发；如果不一致即RPF检查失败，表明该报文来源路径错误，就将其丢弃。

如图12-44所示，来自组播源152.10.2.2的组播流从S1口到达路由器。路由器检查IP 路由表，发现连接152.10.0.0/16 网络的接口是 S0（表示可以转发该组播流的端口为S0），于是RPF检查失败，达到S1口的数据流被丢弃。

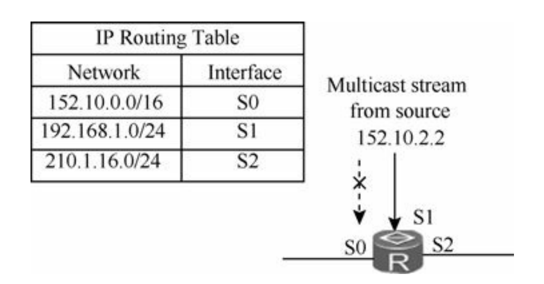


图12-44 RPF检查失败的示例

如果来自组播源 152.10.2.2 的组播流是从S0口达到路由器，则在检查IP路由表时发现入接口与接收该组播流的接口S0一致，RPF检查成功，组播流将被正确地转发。

2．RPF检查在组播数据转发中的应用

组播路由协议通过已有的单播路由、MBGP路由或组播静态路由信息来确定上、下游邻居设备，创建组播路由表项。运用RPF检查机制来确保 组播数据流能够沿组播分发树（路径）正确地传输，同时可以避免转发路径上环路的产生。

在实际组播数据转发过程中，如果对每一份接收到的组播数据报文都通过单播路由表进行RPF检查会给路由器带来很大负担。因此，路由器在收到一份来自源S发往组G的组播数据报文之后，首先会在组播转发表中查找有无相应的（S，G）组播转发表项。

（1）如果不存在（S，G）转发表项，则对该报文执行 RPF 检查，检查通过后将检查到的RPF接口作为入接口，创建组播路由表项，下发到组播转发表中，这样下次来自相同组播组的报文可以不再进行RPF检查。

（2）如果存在（S，G）转发表项，且接收该报文的接口与对应的组播转发表项的入接口一致，则向所有的出接口转发该报文。

（3）如果存在（S，G）转发表项，但是接收该报文的接口与对应的组播转发表项的入接口不一致，则对此报文进行RPF检查。对RPF检查结果的处理方式如下。

① 如果经过 RPF检查得出的 RPF接口与对应的组播转发表项的入接口一致，则说明在组播转发表中的（S，G）表项是正确的，但报文的来源路径有错误，将其丢弃。

② 若RPF检查选取出的RPF接口与对应的组播转发表项的入接口不符，则说明在组播转发表中的（S，G）表项已过时，于是把对应的组播转发表项中的入接口更新为RPF接口。然后根据RPF检查规则进行判

断：如果接收该报文的接口正是其RPF接口，则向转发表项的所有出接口转发该报文，否则将其丢弃。

12.6.3 组播静态路由

前面说了，组播静态路由是RPF检查的重要依据之一。根据不同的应用场景，组播静态路由有如下两种作用。但组播静态路由仅在配置的组播路由器上生效，不会引入或广播给网络中的其他路由器。

1．改变RPF路由

在相同拓扑的网络中可以通过配置组播静态路由改变RPF路由，为组播数据创建一条与单播不同的传输路径。

在如图12-45所示的网络中，RouterC到组播源（Source）的RPF邻居为RouterA， 从 Source 发出的组播报文会沿着Source→RouterA→RouterC 的路径传输。此时，如果在RouterC上配置组播静态路由，指定RouterC的RPF邻居为RouterB（也就相当于“下一跳”），则从 Source 发出的组播报文的传输路径将发生改变，改为沿 Source→RouterA→RouterB→RouterC的路径传输，区别于原来的单播路由路径，可以实现单播和组播报文的分流。

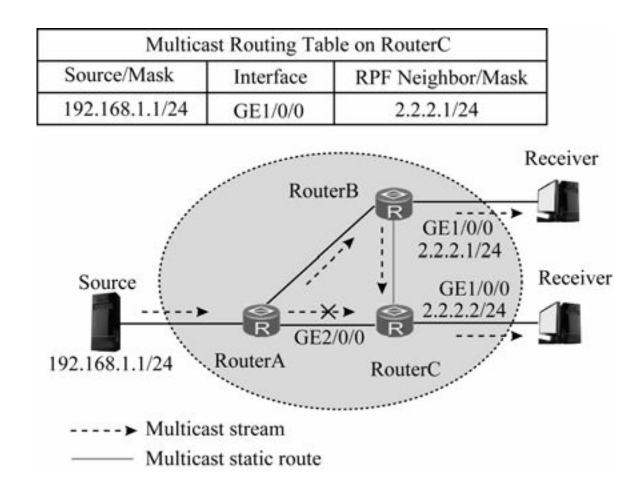


图12-45 配置组播静态路由改变RPF路由示例

2．衔接RPF路由

当某组播网络中的单播路由被阻断时，组播数据流传输将因为没有RPF 路由而中断。此时，可以通过配置组播静态路由，生成新的RPF路由，从而在设备上创建新的组播转发表项来指导组播数据的转发。

如图12-46所示，Domain1和Domain2是单播路由隔离的两个路由域（如RIP和OSPF）， Domain2中的Receiver无法接收来自组播源的组播数据。此时在Domain2中的RouterC、RouterD上分别配置组播静态路由，重新指定其RPF邻居（RouterC邻居指定为RouterB， RouterD邻居指定为RouterC），这样Receiver就能接收组播源的数据流了。从中可以看出，通过配置组播静态路由可以绕过单播路由中路由不通的问题，独立地建立组播传输路径。

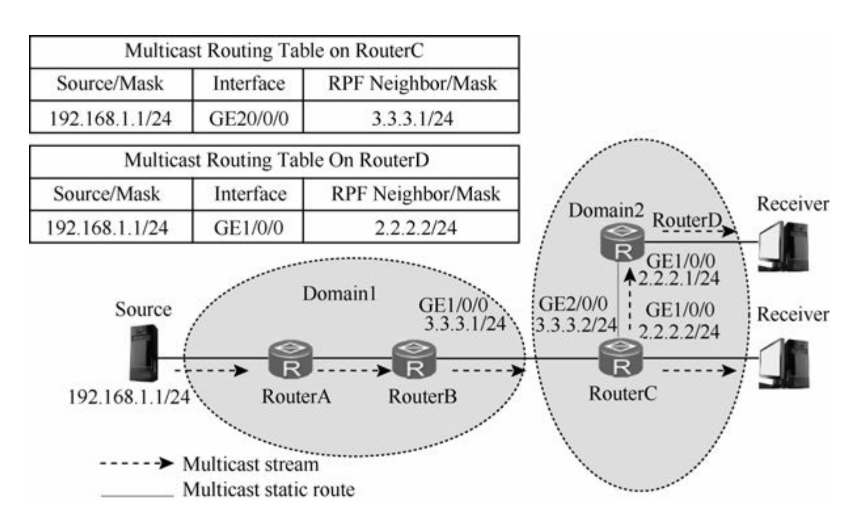


图12-46 配置组播静态路由衔接RPF路由示例

12.6.4 组播负载分担

“负载分担”是指如果发往某一目的地的数据流存在多条等价的转发路径，就将数据在这多条路径上转发，以达到分流的目的。在进行数据转发时，每一条路径上转发的数据流量并不一定相同，转发流量多少需要根据负载分担方式来决定。

缺省情况下，组播报文转发过程中如果存在多条等价的最优转发路径，按照RPF检查对等价路由的处理规则分两种情况。

（1）如果这几条等价路由都是来自同一个路由表，比如单播路由表、组播静态路由表或者MBGP路由表中的一种，则选取下一跳IP地址最大的路由作为RPF路由。

（2）如果这几条等价路由来自不同的路由表，首先会比较路由优先级，再比较掩码长度（掩码长度越长越优先，因为这样的路由更精确）选择一条路由作为RPF路由。如果上述都相同，则设备会根据一定的函数计算选取出一条路由作为RPF路由。

从上分析可以看出，缺省情况下，无论上述哪种情况，设备在RPF检查时都只会选取出一条路由作为RPF路由，进行组播报文转发。

要实现组播负载分担，就需要在多条等价的最优转发路径上根据一定的负载分担方式同时转发流量，不按照以上介绍的RPF检查规则来选取一条最佳的RPF路由。

如图12-47左图所示，组播源Source向组播组G发送组播流，路由器RouterA和RouterD 之间运行某种IGP 协议（如 OSPF ）， RouterA→RouterB→RouterD 和RouterA→RouterC→RouterD是两条等价转发路径。缺省情况下，根据RPF检查规则， 组播流会从Int0端口转发，因为Int0接口的IP地址比Int1接口IP地址大。配置组播负载分担之后，就不会根据下一跳地址来选取转发路径，此时RouterA→RouterB→RouterD和RouterA→RouterC→RouterD这两条路径都会转发组播流，如图12-47右图所示。

对于来自任意源组播（*，G）或指定源（S，G）组播的数据流，“组播负载分担”提供了不同的负载分担方式来支持不同应用场合。

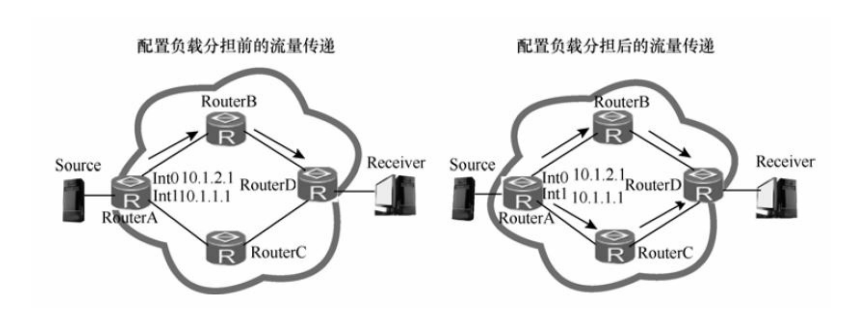


图12-47 配置组播负载分担前后对比示意图

1．基于组播组G的负载分担

如图12-48所示，从同一源Source发往不同组播组 G（G1～G10）的数据流，沿途的 Router7、Router6和Router5 分别存在两条来自源 Source的等价路由。组播路由器经过一系列算法，从等价路径中为不同的组播组G选择一条合适的路径作为转发路由。实现负载分担后，不同转发路径上的流量属于不同的组播组G。

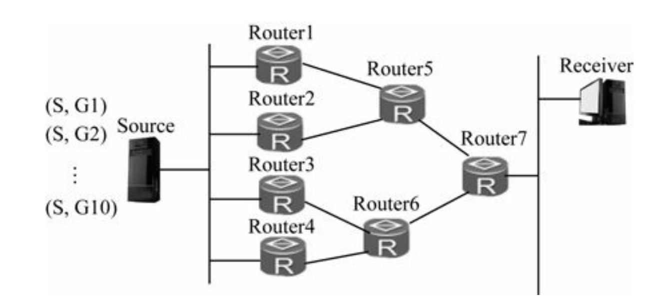


图12-48 基于组播组G的负载分担示意图

2．基于组播源S的负载分担

如图12-49所示，从不同源Source（S1～S10）发往相同组播组G的数据流，沿途的Router7、Router6和Router5也分别存在两条来自源Source的等价路由。组播路由器经过一系列算法，从等价路径中为不同的组播源S选择一条合适的路径作为转发路由。实现负载分担后，不同转发路径上的流量属于不同的组播源S。

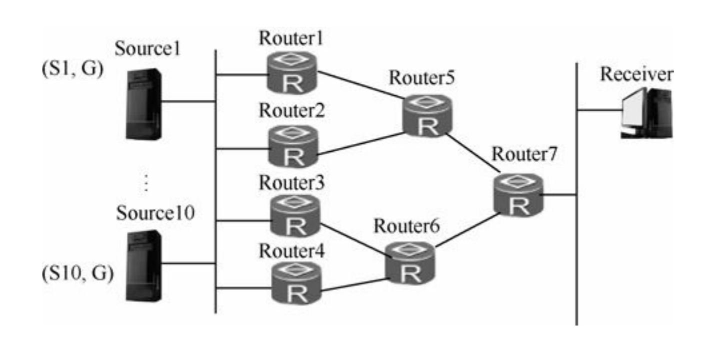


图12-49 基于组播源S的负载分担示意图

3．基于组播源组（S，G）的负载分担

如图12-50所示，从不同源Source（S1～S10）发往不同组播组G（G1～G10）的数据流，沿途的Router7、Router6和Router5也分别存在两条来自源Source的等价路由。组播路由器经过一系列算法，从等价路径中为不同的组播源组（S，G）选择一条合适路径作为转发路由。实现负载分担后，不同转发路径上的

流量属于不同的组播源组（S，G）。

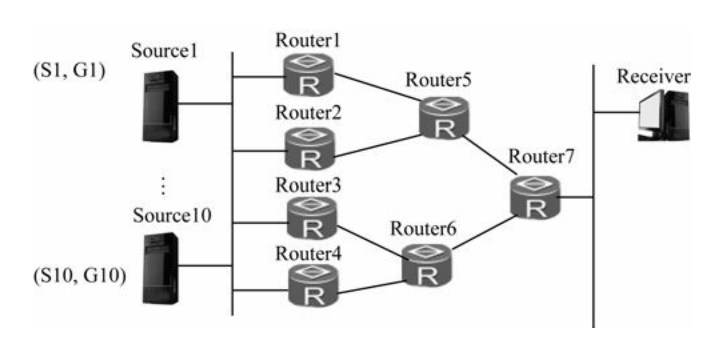


图12-50 基于组播源组（S，G）的负载分担示意图

4．其他负载分担方式

除了以上介绍的几种负载分担方式，还有以下几种负载分担方式。

（1）稳定优先负载分担。如图12-54所示，当组播网络中发生路由振荡时，如果组播路由器频繁调整负载会加剧路由的不断振荡。配置稳定优先负载分担的路由器不会立刻调整负载，而是等到振荡结束后才进行调整。当网络拓扑稳定无振荡时，路由器上来自同一个组播源的路由表项会均衡分布在各等价路径上。

（2）均衡优先负载分担。同样参见图 12-54，当组播网络中发生路由振荡时，配置均衡优先负载分担的路由器，会立刻重新调整负载到均衡状态。当网络拓扑稳定无振荡时，路由器上来自同一个组播源的路由表项会均衡分布在各等价路径上。

（3）不均衡负载分担。同样参见图 12-54，不均衡负载分担是对上述两种“稳定优先”和“均衡优先”负载分担方式的补充，不改变这两种方式的基本行为，只是让路由表项按比例分布在各等价路径上。实际网络中各路径的负载能力存在差异，或者需要人为干预某路径上的负载。不均衡负载分担模式允许在路由器指定接口上配置权值，权值越大的接口所在路径上分布的路由表项越多，从而解决上述问题。

第13章　IP组播配置与管理

13.1 IGMP配置与管理

13.2 PIM-DM（IPv4）配置与管理

13.3 PIM-SM（IPv4）配置与管理

13.4 IGMP Snooping配置与管理

13.5 组播VLAN配置与管理

第12章比较全面地介绍了IP组播相关的基础知识，以及各种组播协议的功能及实现原理，本章要正式介绍各组播协议的各主要功能配置与管理方法。首先要说明的是，本章所说的“组播路由器”也是对运行组播路由协议的三层交换机和路由器的总称。

一个完整的组播网络，一般不是仅包括一种组播协议，而是至少包括像IGMP/MLD、PIM协议这样的三层组播协议，像IGMP Snooping/MLD Snooping这样的二层组播协议。其中IGMP/MLD协议用于三层交换机上的组成员的维护与管理，PIM用于组播路由，IGMP Snooping/MLD Snooping用于二层交换机上的组成员维护和管理。所以在配置一个组播网络时，一定要根据对应的组播网络设备类型完整地配置各设备上运行的对应组播协议功能。

当然，在一个特定的组播网络中，也不是每个组播协议的每个功能都需要配置，也要根据实际需要选择。另外，像MSDP、IGMP Snooping Proxy/MLD Snooping Proxy、IGMP SSM Maping、IGMP Snooping SSM Maping、组播VLAN等这类协议或功能基本上都是可选的，不是必须要配置的。本章仅介绍IPv4网络中的各组播协议功能配置，但均不考虑多VPN实例情形，且由于篇幅限制，不介绍用于多PIM域组播的MSDP配置。

13.1 IGMP配置与管理

IGMP（Internet Group Management Protocol，因特网组管理协议）是TCP/IP协议族中负责IPv4组播成员管理的协议，需要在组播组成员主机和与之相连的组播路由器上运行，用来在组播组成员主机和与其直接相邻的组播路由器之间建立、维护组播组成员关系。

到目前为止，IGMP有3个版本：IGMPv1版本（由RFC1112定义）、IGMPv2版本（由RFC2236定义）和IGMPv3（由RFC3376定义）版本。所有IGMP版本都支持ASM （任意源组播）。运行IGMPv3的主机可以直接应用于SSM（指定源组播）模型，而运行IGMPv1和 IGMPv2的主机则需要与在 IGMP交换机上运行的SSM Mapping结合才能应用于SSM模型中。

说明

在本节后面的配置中，凡同时支持全局配置（即IGMP配置）和接口配置，则最终的配置生效原则如下。

（1）在IGMP视图下的配置全局有效，在接口视图下的配置只对该接口有效。

（2）如果接口视图和IGMP视图下都配置了命令，则优先选择接口视图下配置的值。接口视图下没有配置时，IGMP视图下配置的值有效。

（3）如果IGMP视图下配置非缺省值，则接口视图下配置的缺省值无效。

13.1.1 IGMP特性的产品支持

在交换机上配置 IGMP 协议时，可以配置的功能特性包括 IGMP 基本功能、IGMP性能调整（包括Router-Alert选项、IGMP查询控制器、快速离开和IGMP报文过滤等）、SSM-Mapping、IGMP组成员关系个数限制等。

注意

但这里要特别说明的是，IGMP是三层组播协议，需要运行在配置了IP地址的三层接口上，但华为S系列交换机的物理以太网端口是不能直接配置IP地址的，所以在配置IGMP协议时只能选择可以配置IP地址的VLANIF接口或者Loopback接口（在S9300/9300E/9700系列中还支持 POS接口和IP-Trunk接口）。在本章配置中，如无特殊说明，接口的配置一般选择 VLANIF接口，但需要事先将对应的物理接口加入到该VLAN中。

1．IGMP基本功能

华为S系列交换机中的IGMP基本功能配置包括以下几个方面。

（1）IGMP版本配置：支持IGMPv1、IGMPv2和IGMPv3，版本可配置。由于不同版本的IGMP协议报文不相同，因此，需要为交换机和组播组成员主机配置匹配的版本（交换机侧的高版本可以兼容主机侧的低版本）。现在一般的主机操作系统都支持IGMPv3了，如Windows XP、Windows Server 2003、Windows Vista、Windows 7和Windows 8等。

（2）配置静态加入组播组：当网络中存在稳定的组播组成员时，通过配置交换机用 户侧接口静态加入指定的组播组，可以实现组播数据的快速、稳定转发。

（3）配置接口允许加入的组播组范围：通过在交换机用户侧对应接口上设置一个ACL规则作为过滤器，就可以限制该接口所服务的组播组范围，从而控制组播数据的安全发送。

2．IGMP性能调整

在IGMP性能调整方面，可配置以下功能。

（1）Router-Alert选项：配置设备仅接收包含Router-Alert选项的IGMP报文，提高安全性。

（2）查询器：对IGMP查询器的参数进行合理配置，既可以使成员关系得到及时的更新维护，又可以避免报文发送过多造成网络拥塞。

（3）快速离开：使IGMP协议可以快速响应成员主机的Leave（离开）报文。

（4）IGMP On-Demand：可根据组播成员的实际需求维护组成员之间的关系，减少了报文交互，降低网络流量。

（5）IGMP报文过滤：可根据组播报文中源IP地址来过滤用户侧IGMP接口收到的IGMP报文，提高安全性。

3．SSM Mapping

SSM是一种在IGMPv3协议支持的情况下，能够在用户侧IGMP接口上指定组播源的传输服务。但有时组播组成员主机却只能运行IGMPv1或IGMPv2，这时就可以通过在交换机上配置SSM Mapping功能，向运行 IGMPv1或 IGMPv2的组成员提供SSM服务。

4．IGMP Limit

IGMP协议规定，成员可以在任意时间、任意位置、成员总数不受限制地加入或退出组播组。但是当大量用户同时收看多套节目时，需要占用组播设备的大量带宽，可能会造成组播性能下降。为了避免这种情况的发生，交换机支持 IGMP Limit功能，通过限制全局和用户侧IGMP接口下的组播组个数，使加入组播组的用户收看更加清晰稳定的节目。

在以上四方面的特性中，S2700/3700系列交换机除不支持“IGMP报文过滤”和“IGMP Limit”特性外，其他特性均支持，其他系列均支持上述所有IGMP特性。但各系列缺省情况下可支持的并发用户数不同：

S5700EI系列最多能够同时处理大约150个组播用户的点播需求；S5710EI/5700HI/7700/9300/9300E/9700系列最多能够同时处理大约190个组播用户的点播需求；S6700系列最多能够同时处理大约290个组播用户的点播需求。

13.1.2 配置IGMP基本功能

在成员主机和与之相连的交换机上配置 IGMP，在此仅介绍在交换机上配置 IGMP的方法。在以下介绍的IGMP配置任务中，仅“使能IGMP功能”和“配置IGMP版本”为必选配置，其他为可选配置，请根据需要选配。

说明

缺省情况下，华为S系列交换机未使能 IP组播路由、IGMP和 IGMP SSM Mapping 功能，IGMP的版本为IGMPv2。

1．IGMP配置任务

通过在与用户网段相连的组播设备（此处指**S**系列交换机，下同 ）接口上使能IGMP 基本功能，用户主机可以接入组播网络，接收组播报文。但在配置IGMP基本功能之前，需配置单播路由协议，使各节点间IP路由可达。主要的配置任务如下。

（1）使能IGMP功能。配置IGMP协议之前，必须先使能IP组播路由功能，因为它是配置一切组播功能的前提。如果停止 IP 组播路由，组播所有相关配置将无法生效。**IGMP**应该配置在与组成员相连的接口上。

（2）配置IGMP版本。运行IGMP高版本的交换机可以识别低版本的成员报告，但是低版本的交换机不能识别高版本的成员报告。为了保证IGMP的正常运行，建议在交换机上配置与组播组成员主机上运行相同，或高于组播组成员主机的版本。

如果在主机侧共享网段上有多个交换机，由于不同版本的IGMP协议报文结构不同，为了保证IGMP的正常运行，必须在所有交换机接口配置相同的IGMP版本。

此项配置同时支持全局配置（即**IGMP**视图）和接口配置，生效原则参见本节前面说明。

（3）（可选）配置静态组播组。在以下应用场景中，可在交换机的用户侧接口上配置静态组播组。

① 网络中存在稳定的组播组成员。

② 某网段内没有组播组成员或组播组成员主机无法发送Report报文，但是又需要将组播数据转发到该网段，可以在接口上配置静态组播组，将组播数据“拉”到接口上。

在接口上配置静态组播组后，交换机就认为此接口网段上一直存在该组播组的成员，从而转发该组的组播数据。

（4）（可选）配置接口加入的组播组范围。为了让 IGMP 接口所在网段的组播组成员主机加入指定的组播组，并接收这些组的报文，可以在该接口上设置ACL规则，对收到的成员 Report 报文进行过滤，使交换机只对该规则中允许的组播组维护组成员关系。

2．配置步骤

以上四大IGMP基本功能配置任务的具体配置步骤如表13-1所示。

表13-1 IGMP基本功能配置步骤

配置任务	步骤	命令	说明
公共 配置步骤	1	**system-view** 例如：<HUAWEI> **system-view**	进入系统视图
使能 IGMP 功能	2	**multicast routing-enable** 例如：[HUAWEI] **multicast routing-enable**	使能 IP 组播路由功能。全局使能组播路由功能是配置三层组播功能的前提，即只有在使能了组播路由功能之后，才能配置 PIM、IGMP 等一些三层组播协议以及其他三层组播功能 缺省情况下，没有使能组播路由功能，可用 **undo multicast routing-enable** 命令去使能组播路由功能 【注意】使用 **undo multicast routing-enable** 命令将清除设备上所有的组播配置。如果设备上正在运行组播业务，则组播业务将会中止；如果下次需要恢复组播业务，必须重新配置被清除掉的组播命令

（续表）

配置任务	步骤	命令	说明	
	3	**interface** *interface-type interface-number* 例如：[HUAWEI] **interface vlanif** 10	键入要使能 IGMP 功能的 VLANIF 或者 Loopback 接口（**不能直接键入物理接口，但要把对应的物理接口加入此 VLAN 中**），进入对应的接口视图	
	4	**igmp enable** 例如：[HUAWEI-Vlanif10] **igmp enable**	在以上 VLANIF 或者 Loopback 接口上使能 IGMP 功能，这样组播设备才能处理来自主机的协议报文 【注意】如果接口上需要同时使能 PIM 和 IGMP，必须要先使能 PIM，**再使能** IGMP；使能 IGMP 前如果接口上配置了其他 IGMP 参数，只有在配置了此命令后才生效	
配置 IGMP 版本	5	**quit** 例如：[HUAWEI-Vlanif10] **quit**	退出接口视图，返回系统视图	
	6	**igmp** 例如：[HUAWEI] **igmp**	进入 IGMP 视图。与 IGMP 相关的全局参数必须在 IGMP 视图下配置 可用 **undo igmp** 命令清除 IGMP 视图下的所有配置	配置 全局的 IGMP 版本
	7	**version** { **1** | **2** | **3** } 例如：[HUAWEI-igmp] **version** 3	在全局上配置 IGMP 的版本，所配置的版本将应用于本地交换机上所有使能了 IGMP 功能的接口。为了保证正常工作，需要在同网段所有组播设备上配置相同版本的 IGMP，因为 IGMP 各版本之间不能自动转换 缺省情况下，IGMP 的版本是 IGMPv2，可用 **undo version** 命令恢复缺省的 IGMPv2 版本。**如果此处配置的是非缺省值，则下面在接口视图下配置的缺省值无效**	
	8	**interface** *interface-type interface-number* 例如：[HUAWEI] **interface vlanif** 10	（可选）再次键入前面使能了 IGMP 功能的 VLANIF 或者 Loopback 接口，进入对应的接口视图	配置 接口的 IGMP 版本
	9	**igmp version** { **1** | **2** | **3** } 例如：[HUAWEI-Vlanif10] **igmp version** 3	（可选）在以上 VLANIF 或者 Loopback 接口上配置 IGMP 版本，仅作用于此接口 缺省情况下，接口上运行 IGMPv2，可用 **undo igmp version** 命令恢复为缺省的 IGMPv2 版本	
（可选） 配置静态 组播组	10	**igmp static-group** *group-address* [**inc-step-mask** { *group-mask* | *group-mask-length* } **number** *group-number*] [**source** *source-address*] 例如：[HUAWEI-Vlanif10] **igmp static-group** 225.1.1.1 **inc-step-mask** 32 **number** 10	在以上 VLANIF 或者 Loopback 接口上配置静态组播组，使该接口静态加入一个或者一个范围的组播组。命令中的参数说明和选项如下。 （1）*group-address*：指定接口要加入的组播组 IP 地址，为 D 类组播地址，取值范围是 224.0.1.0～239.255.255.255。如果为批量配置方式，则为组地址序列的起始组播组地址 （2）**inc-step-mask**：可选项，指定批量配置方式中的各组播组地址间的递增掩码。可以通过下面的 *group-mask*（组播组地址递增掩码）或者 *group-mask-length*（组播组地址递增掩码长度）来表示 （3）*group-mask*：二选一可选参数，指定批量配置方式中的组播组地址递增掩码，即组播组地址序列中相邻两个组播组地址的间隔。**它采用反掩码**（即	

（续表）

配置任务	步骤	命令	说明
（可选）配置静态组播组	10	**igmp static-group** *group-address* [**inc-step-mask** { *group-mask* | *group-mask-length* } **number** *group-number*] [**source** *source-address*] 例如：[HUAWEI-Vlanif10] **igmp static-group** 225.1.1.1 **inc-step-mask** 32 **number** 10	**子网掩码的反码）形式表示**，点分十进制形式，取值范围是 0.0.0.1～255.255.255.255，用于表示一个组播组地址范围 （4）*group-mask-length*：二选一可选参数，指定批量配置方式中的组播组地址递增掩码长度（**值为 1 的连续位长度**），取值范围为 4～32 的整数，也可用于表示一个组播组地址范围。**当组播组地址步长掩码长度为 32 时表示任意组播组地址**。但使用本参数配置组播组地址递增掩码后，在使用 **display current-configuration** 命令查看相关配置信息时，显示的组地址递增范围掩码仍将转换为 *group-mask* 格式 （5）**number** *group-number*：可选参数，指定批量配置方式中接口可加入的组播组地址个数，取值范围为 2～512 的整数 （6）**source** *source-address*：可选参数，指定允许静态加入的组播组中的组播源 IP 地址（是一个单播 IP 地址），此时接口中的组播转发表中为 SSM 模式的（S，G）格式，如果不指定此可选参数，则为 ASM 模型的（*，G）格式 【注意】执行本命令后，接口上的 IGMP 静态组记录永远不会超时，交换机会认为该接口上始终连接着组成员主机，并持续向该接口所在网段转发符合条件的组播报文。当组播组成员不再需要静态加入的组播组数据时，需要手动删除静态组播组配置。不同的组播组批量配置可以存在相同的组播组地址第一次配置批量组播组后，若再配置批量组播组时只修改 *group-number* 的配置值，不改变 *group-address* 和 *group-mask* | *group-mask-length* 的配置值，则会覆盖之前的批量组播静态组配置 缺省情况下，接口未配置任何静态组播组，可用 **undo igmp static-group** { **all** | *group-address* [**inc-step-mask** { *group-mask* | *group-mask-length* } **number** *group-number*] [**source** *source-address*] }命令删除接口上配置的静态组播组
（可选）配置接口加入的组播组范围	11	**igmp group-policy** { *acl-number* | **acl-name** *acl-name* } [**1** | **2** | **3**] 例如：[HUAWEI-Vlanif10] **igmp group-policy** 2001 2	在接口上设置 IGMP 组播组的过滤策略，限制组播组成员主机能够动态加入的组播组范围。命令中的参数和选项说明如下。 （1）{ *acl-number* | **acl-name** *acl-name* }：指定要在策略中用于过滤成员主机发送的 Report 报文的数字型或者命名型 ACL。在定义 ACL 的 rule 时，通过 **permit** 参数仅允许接口下成员主机可以加入指定地址范围的组播组，如果指定的 ACL 未定义规则，则禁止接口下成员主机加入所有组播组。有关 ACL 方面的知识具体参见本书第 9 章 （2）[**1** | **2** | **3**]：可选项，指定要通过 ACL 限制加入特定组播组的组成员主机运行的 IGMP 版本。如果不指定 IGMP 版本，则该 ACL 同时适用于 IGMPv1、v2 和 v3 版本的主机

（续表）

配置任务	步骤	命令	说明
（可选）配置接口加入的组播组范围	11	**igmp group-policy** { *acl-number* | **acl-name** *acl-name* } [**1** | **2** | **3**] 例如：[HUAWEI-Vlanif10] **igmp group-policy** 2001 2	【说明】为了让接口所连接网络上的主机加入指定范围的组播组，并接收这些组的报文，可以使用本命令在对应接口上设置一个 ACL 规则作为过滤器，以限制接口所服务的组播组范围，从而提高 IGMP 的安全性。当交换机不希望接收某些组的加入报文，不希望转发该组播组的数据时，也可以通过本命令加以限制 缺省情况下，接口可以加入任何组播组，可用 **undo igmp group-policy** 命令取消接口上配置的组播组过滤策略

【示例 1】在与用户相连的VLANIF100接口（先要把对应物理接口加入VLAN 100中）配置静态组播组224.1.1.1。

<HUAWEI>system-view

[HUAWEI] multicast routing-enable

[HUAWEI] interface vlanif 100

[HUAWEI-Vlanif100] igmp static-group 224.1.1.1

【示例 2】配置VLANIF100接口加入组播源IP地址为192.168.10.1，组播组IP地址为232.1.1.1的组播组（192.168.10.1，232.1.1.1）中。

<HUAWEI> system-view

[HUAWEI] multicast routing-enable

[HUAWEI] interface vlanif 100

[HUAWEI-Vlanif100] igmp static-group 232.1.1.1 source 192.168.10.1

【示例 3】配置VLANIF100接口加入以225.1.1.1为起始组播组IP地址，组播组地址递增掩码长度为8（相当于组播组子网碗码为255.255.255.0），组播组地址数量限制为10的批量组播组中。

<HUAWEI> system-view

[HUAWEI] multicast routing-enable

[HUAWEI] interface vlanif 100

[HUAWEI-Vlanif100] igmp static-group 225.1.1.1 inc-step-mask8number 10

【示例 4】配置VLANIF100接口加入以232.1.1.1为起始组播组IP地址，组播源IP地址为192.168.11.1，组播组地址递增掩码为0.0.255.255，组地址数量限制为10的批量组播组中。

<HUAWEI>system-view

[HUAWEI] multicast routing-enable

[HUAWEI] interface vlanif 100

[HUAWEI-Vlanif100] igmp static-group 232.1.1.1 inc-step-mask0.0.255.255 number10 source192.168.11.1

【示例 5】创建编号为2005的ACL，允许主机接收来自组播组225.1.1.1的数据，然后在VLANIF100接口应用过滤策略，以限定该接口下的主机只能加入组播组225.1.1.1。

<HUAWEI> system-view

[HUAWEI] acl number 2005

[HUAWEI-acl-basic-2005] rule permit source 225.1.1.1 0

[HUAWEI-acl-basic-2005] quit

[HUAWEI] multicast routing-enable

[HUAWEI] interfacevlanif 100

[HUAWEI-Vlanif100] igmp group-policy 2005

13.1.3 调整IGMP性能

使能IGMP后，缺省情况下可以正常工作。但也可根据安全性和网络性能优化的要求适当调整相关参数。当然，事先要配置好上节介绍的IGMP基本功能，否则所配置的参数不会立即生效。可以调整的IGMP性能参数包括以下几个方面（均为可选配置任务 ）。

1．配置Router-Alert选项

通常情况下，网络设备收到报文时，只有目的IP地址为本设备接口地址的报文才会上送给相应的协议模块处理。这样就会存在一个问题，如果协议报文的目的地址不为本设备的接口地址，比如IGMP协议报文，由于其目的地址为组播地址，这种情况下就无法上送给IGMP协议模块处理，导致正常的组成员关系不能维护。为了解决此类问题， Router-Alert选项应运而生。如果IP报文头中携带Router-Alert选项，设备在接

收到此类报文后会直接上送给相应的协议模块处理，而不检查目的地址。

缺省情况下，出于兼容性考虑，当前交换机在收到IGMP报文后，无论其IP报文头是否包含Router-Alert选项都会上送给IGMP协议模块处理。交换机在发送IGMP报文时，也可以选择是否需要携带 Router-Alert 选项。缺省情况下，组播设备发送的 IGMP报文中携带Router-Alert选项。

此项配置同时支持全局配置（即**IGMP**视图）和接口配置，生效原则参见本节前面说明。

2．配置IGMP查询器参数

IGMP通过查询/响应报文维护组成员关系。当同一网段 上有多台组播设备时，是由IGMP查询器负责发送IGMP查询报文，这时就需要指定IGMP查询器（在**IGMPv1**中，查询器是由**PIM**协议指定的，**IGMPv2**和**IGMPv3**可以手动配置 ）。IGMP查询器在工作过程中使用了表13-2所示的多项参数，缺省情况下这些参数可以正常工作。同时根据需要，也可以通过命令行进行调整。

表13-2 可以调整的IGMP性能参数

查询器参数	参数说明	支持的版本
IGMP 普遍组查询报文的发送时间间隔	查询器周期性地发送普遍组查询报文，维护接口上的组成员关系，本参数定义了发送该报文的时间间隔（**缺省值为 60s**）	IGMPv1、IGMPv2、IGMPv3
IGMP 健壮系数	健壮系数是指用来弥补可能发生的网络丢包而设置的消息重传次数（**缺省值为 2**）用来规定以下两个值。 （1）当 IGMP 查询器启动时发送“健壮系数”次的“普遍组查询报文”，发送时间间隔为“IGMP 普遍组查询报文发送间隔”的 1/4 （2）当组播设备收到 Leave 报文后，发送“健壮系数”次的“IGMP 特定组查询报文”，发送间隔为“IGMP 特定组查询报文发送间隔”	IGMPv1、IGMPv2、IGMPv3
IGMP 查询报文的最大响应时间	组播组成员接收到一个 IGMP 查询报文后，会在最大响应时间（**缺省值为 10s**）内发送 Report 报文	IGMPv2、IGMPv3
其他 IGMP 查询器的存活时间	如果非查询器在“其他 IGMP 查询器的存活时间”内收不到查询报文，就认为查询器失效，自动发起查询器选举 “其他 IGMP 查询器存活时间”=“普遍组查询报文发送间隔”*“健壮系数”+“最大响应时间”*（1/2）。当等式右边的参数都取缺省值时，“其他 IGMP 查询器存活时间”的值为 125s	IGMPv2、IGMPv3
IGMP 特定组查询报文的发送间隔	当查询器收到主机退出某组播组的 Leave 报文时，会连续发送特定组查询报文，询问该组播组是否还存在成员。本参数定义了发送该报文的时间间隔（**缺省值为 1s**）	IGMPv2、IGMPv3

在实际配置中，要确保“IGMP查询报文最大响应时间”<“IGMP普遍组查询报文发送间隔”<“其他IGMP查询器存活时间”。在共享网段内，如果多台设备的用户侧接口都使能了IGMP，应确保设备上配置的查询器参数一致，否则有可能导致IGMP协议无法正常运行。

此项配置同时支持全局配置（即**IGMP**视图）和接口配置，生效原则参见本节前面说明。

3．配置IGMP快速离开

在某些应用中，IGMP查询器的一个接口下只连接着一台成员主机（这是前提条件 ），当主机需要在多个组播组间频繁切换时，为了快速响应主机的离开组报文，可以在IGMP查询器上配置 IGMP 快速离开功能。这样，当查询器收到来自主机的 Leave（离开）报文时，不再发送特定组查询报文，而是直接向上游发送离开通告。这样可减小响应延迟，也节省网络带宽。

IGMP快速离开功能仅适用于**IGMPv2**和**IGMPv3**版本。此项配置同时支持全局配置（即**IGMP**视图）和接口配置，生效原则参见本节前面说明。

4．配置 IGMP On-Demand

在标准的 IGMP 工作机制中，查询器通过周期性发送查询报文并接收成员反馈的Report和Leave报文来了解组播组成员信息，组成员收到查询时都会进行回应。为了减少这个过程中的报文交互，降低网络流量，可以在查询器上配置 IGMP On-Demand功能。使能了 IGMP On-Demand功能后，查询器可根据组播组成员的要求来维护成员关系，不再主动发送查询报文来收集成员状态。

IGMP On-Demand只适用于 **IGMPv2**和 **IGMPv3**。

5．配置根据源地址过滤IGMP报文

为了提高安全性，可以在交换机的接口上对IGMP报文（包括Query报文、Report和Leave报文）进行过滤。

除**S2700/3700**系列交换机外，其他所有**S**系列交换机均支持**IGMP**报文过滤。

以上五方面的IGMP性能参数配置步骤如表13-3所示。

表13-3 调整IGMP性能参数的配置步骤

配置任务	步骤	命令	说明	
公共配置步骤	1	**system-view** 例如：<HUAWEI> **system-view**	进入系统视图	
配置Router-Alert 选项	2	**igmp** 例如：[HUAWEI] **igmp**	进入 IGMP 视图，其他说明参见表 13-1 中的第 2 步	配置全局 Router-Alert 选项
	3	**require-router-alert** 例如：[HUAWEI-igmp] **require-router-alert**	全局配置丢弃 IP 报文头中不包含 Router-Alert 选项的 IGMP 消息。这样在当交换机接收到 IGMP 消息检查该 IP 报文头中的 Router-Alert 选项时，如果不包含该选项，就丢弃这个 IGMP 消息 缺省情况下，交换机不对 Router-Alert 选项进行检查，即处理所有接收到的 IGMP 消息，可用 **undo require-router-alert** 命令全局恢复缺省配置	

（续表）

配置任务	步骤	命令	说明	
	4	**send-router-alert** 例如：[HUAWEI-igmp] **send-router-alert**	全局指定该交换机发送的 IGMP 报文的 IP 报头中包含 Router-Alert 选项 缺省情况下，该交换机发送的 IGMP 报文头中包含 Router-Alert 选项，可用 **undo send-router-alert** 命令全局指定该交换机发送的 IGMP 报文的报文头中不包含 Router-Alert 选项	
	5	**interface** *interface-type interface-number* 例如：[HUAWEI] **interface vlanif** 10	（可选）键入前面使能了 IGMP 功能的 VLANIF 或者 Loopback 接口，进入接口视图	
	6	**igmp require-router-alert** 例如：[HUAWEI-Vlanif10] **igmp require-router-alert**	（可选）在以上接口上配置丢弃 IP 报文头中不包含 Router-Alert 选项的 IGMP 消息 缺省情况下，接口不对 Router-Alert 选项进行检查，即处理所有接收到的 IGMP 报文，可用 **undo igmp require-router-alert** 命令恢复接口为缺省配置	配置接口下 Router-Alert 选项
	7	**igmp send-router-alert** 例如：[HUAWEI-Vlanif10] **igmp send-router-alert**	（可选）在以上接口上配置发送的 IGMP 消息其 IP 报文头中包含 Router-Alert 选项 缺省情况下，该接口发送的 IGMP 消息其 IP 报文头中包含 Router-Alert 选项，可用 **undo igmp send-router-alert** 命令在接口上配置发送的 IGMP 消息其 IP 报文头中不包含 Router-Alert 选项	
配置 IGMP 查询器参数	8	**quit** 例如：[HUAWEI-Vlanif10] **quit**	退出接口视图，返回系统视图	
	9	**igmp** 例如：[HUAWEI] **igmp**	进入 IGMP 视图	
	10	**timer query** *interval* 例如：[HUAWEI-igmp] **timer query** 125	全局配置设备发送 IGMP 普遍组查询报文的时间间隔，取值范围为 1～18 000 的整数秒 缺省情况下，IGMP 普遍组查询消息的发送间隔为 60s，可用 **undo timer query** 命令全局恢复该配置参数的缺省值	
	11	**robust-count** *robust-value* 例如：[HUAWEI-igmp] **robust-count 3**	全局配置 IGMP 查询器健壮系数，这是用来弥补可能发生的网络丢包而设置的消息重传次数，取值范围为 2～5 的整数 缺省情况下，IGMP 查询器的健壮系数是 2，可用 **undo robust-count** 命令恢复全局配置为缺省值	配置全局 IGMP 查询器参数
	12	**max-response-time** *interval* 例如：[HUAWEI-igmp] **max-response-time** 15	全局配置 IGMP 查询报文的最大响应时间，取值范围为 1～25 的整数秒。主机响应时间越小，IGMP 设备获知组播成员的速度越快，但是网络带宽和交换机资源的占用也就越大 缺省情况下，IGMP 查询报文的最大响应时间是 10s，可用 **undo max-response-time** 命令全局恢复该配置参数的缺省值	

（续表）

配置任务	步骤	命令	说明	
	13	**timer other-querier-present** *interval* 例如：[HUAWEI-igmp] **timer other-querier-present** 50	全局配置其他 IGMP 查询器的存活时间，取值范围为 60～300 的整数秒。这是用来确定查询器是否有效的时间参数，超时后本地交换机会发起查询器选举 缺省情况下，其他 IGMP 查询器的存活时间的计算公式是：其他 IGMP 查询器的存活时间＝健壮系数×IGMP 普遍查询消息发送间隔+（1/2）×最大查询响应时间。当健壮系数、IGMP 普遍查询消息发送间隔和最大查询响应时间都取缺省值时，其他 IGMP 查询器的存活时间的值为 125s，可用 **undo timer other-querier-present** 命令全局恢复该配置参数的缺省值	
	14	**lastmember-queryinterval** *interval* 例如：[HUAWEI-igmp] **lastmember-queryinterval** 60	全局配置 IGMP 查询器在收到主机发送的 IGMP Leave 报文时，发送 IGMP 指定组查询报文的时间间隔，取值范围为 1～5 的整数秒 缺省情况下，发送 IGMP 指定组查询报文的时间间隔是 1s，可用 **undo lastmember-queryinterval** 命令全局恢复该配置参数的缺省值	
配置 IGMP 查询器参数	15	**quit** 例如：[HUAWEI-igmp] **quit**	退出 IGMP 视图，返回系统视图	配置接口下 IGMP 查询器参数
	16	**interface** *interface-type interface-number* 例如：[HUAWEI] **interface vlanif** 10	（可选）键入要配置 IGMP 查询器参数的 VLANIF 或者 Loopback 接口（当然该接口必须已使能 IGMP），进入接口视图	
	17	**igmp timer query** *interval* 例如：[HUAWEI-Vlanif10] **igmp timer query** 120	（可选）在接口上配置设备发送 IGMP 普遍组查询报文的时间间隔，取值范围为 1～18 000 的整数秒 缺省情况下，IGMP 普遍组查询消息的发送间隔是 60s，可用 **undo igmp timer query** 命令恢复为缺省值	
	18	**igmp robust-count** *robust-value* 例如：[HUAWEI-Vlanif10] **igmp robust-count** 3	（可选）在接口上配置 IGMP 查询器的健壮系数，取值范围为 2～5 的整数 缺省情况下，IGMP 查询器的健壮系数是 2，可用 **undo igmp robust-count** 命令恢复为缺省值	
	19	**igmp max-response-time** *interval* 例如：[HUAWEI-Vlanif10] **igmp max-response-time** 20	（可选）在接口上配置 IGMP 查询报文的最大响应时间，取值范围为 1～25 的整数秒 缺省情况下，IGMP 查询报文的最大响应时间是 10s，可用 **undo igmp max-response-time** 命令恢复接口上该配置参数的缺省值	

（续表）

配置任务	步骤	命令	说明	
配置 IGMP 查询器参数	20	**igmp timer other-querier-present** *interval* 例如：[HUAWEI-Vlanif10] **igmp timer other-querier-present** 100	（可选）在接口上配置其他 IGMP 查询器的存活时间，取值范围为 60～300 的整数秒 缺省情况下的配置与前面的 **timer other-querier-present** 的缺省情况一样，可用 **undo igmp timer other-querier-present** 命令恢复为缺省值	配置接口下 IGMP 查询器参数
	21	**igmp lastmember-queryinterval** *interval* 例如：[HUAWEI-Vlanif10] **igmp lastmember-queryinterval** 3	（可选）在接口上 IGMP 查询器在收到主机发送的 IGMP Leave 报文时，发送 IGMP 最后组成员查询报文的时间间隔，取值范围为 1～5s 缺省情况下，发送 IGMP 最后组成员查询报文的时间间隔是 1s，可用 **undo igmp lastmember-queryinterval** 命令恢复为缺省值	
配置 IGMP 快速离开	22	**quit** 例如：[HUAWEI-Vlanif10] **quit**	退出接口视图，返回系统视图	配置全局的 IGMP 快速离开
	23	**igmp** 例如：[HUAWEI] **igmp**	进入 IGMP 视图	
	24	**prompt-leave** [**group-policy** *acl-number*] 例如：[HUAWEI-igmp] **prompt-leave group-policy** 2010	配置 IGMP 快速离开，当组播设备接收到针对某组播组的离开消息时，不发送最后组成员查询消息，立即删除该组记录。参数 *acl-number* 用来指定要配置离开策略的 ACL 编号，**在 S2700/3700 系列交换机中仅支持基本 ACL**，取值范围为 2 000～2 999 的整数，**其他 S 系列同时支持基本 ACL 和高级 ACL**（取值范围为 2 000～3 999）。如果未配置此参数，则对所有的组播组都执行立即离开 在定义 ACL 规则时，通过 **permit** 参数仅允许接口下成员主机快速离开指定地址范围的组播组。如果 ACL 未定义规则，则禁止接口下成员主机快速离开所有组播组 缺省情况下，IGMP 在接收到主机发送的离开消息后发送最后组成员查询消息，可用 **undo prompt-leave** 命令全局取消快速离开组机制	
	25	**quit** 例如：[HUAWEI-igmp] **quit**	退出 IGMP 视图，返回系统视图	配置接口的 IGMP 快速离开
	26	**interface** *interface-type interface-number* 例如：[HUAWEI] **interface vlanif** 10	（可选）键入要配置 IGMP 查询器参数的 VLAN 或者 Loopback 接口（当然该接口必须已使能 IGMP），进入接口视图	

（续表）

配置任务	步骤	命令	说明
配置 IGMP 快速离开	27	**igmp prompt-leave** [**group-policy***acl-number*] 例如：[HUAWEI-Vlanif10] **igmp prompt-leave group-policy** 2010	（可选）在接口上配置立即离开组。当接口接收到针对某组播组的 Leave 报文时，不发送特定查询报文，立即删除该组记录。其他说明参见上面的第 24 步 缺省情况下，IGMP 查询器在接收到主机发送的 Leave 报文后发送特定组查询报文，可用 **undo igmp prompt- leave** 命令在接口上取消快速离开组机制 配置接口的 IGMP 快速离开
配置 IGMP On-Demand	28	**igmp on-demand** 例如：[HUAWEI-Vlanif10] **igmp on-demand**	（可选）在以上接口上配置 IGMP on-demand 功能，使接口上动态加入的组播组永不超时。使用 **igmp on-demand** 命令后，与IGMP 标准协议行为有3点不同。 （1）接口不发送 IGMP 查询报文 （2）接口上动态加入组播组后，创建的表项永不超时 （3）接口收到 IGMP Leave 报文后，会立即删除接口上相应的 IGMP 组记录 缺省情况下，接口上动态加入的组播组定时老化，可使用 **undo igmp on-demand** 命令恢复缺省配置
配置根据源地址过滤 IGMP 报文	29	**igmp query ip-source-policy** { *basic-acl-number* | **acl-name** *acl-name* } 例如：[HUAWEI-Vlanif10] **igmp query ip-source-policy** 2001	（可选）在以上接口上配置 IGMP Query 报文源地址过滤策略。参数 { *basic-acl-number* | **acl-name** *acl-name* }用来指定用于创建过滤策略的数字型 ACL 或命名型 ACL（**但仅支持基本 ACL**）。在定义 ACL 规则时，通过 **permit** 参数配置接口仅接收指定源地址范围的 Query 报文。如果 ACL 未定义规则，则接口缺省过滤掉所有源地址范围的 Query 报文 【说明】IGMP Query 源地址过滤是一种安全策略，可避免恶意设备伪造 IP 地址相对较小的 IGMP Query 报文，使真正的查询器失效，无法响应组成员快速离开，造成流量浪费。配置此功能后，设备只接收源地址属于 ACL 过滤规则范围内的 IGMP Query 报文，从而控制查询器的选举 缺省情况下，交换机不对 Query 报文进行过滤，即处理所有接收到的 Query 报文，可用 **undo igmp query ip-source-policy** 命令恢复缺省配置
	30	**igmp ip-source-policy** [*basic-acl-number*] 例如：[HUAWEI-Vlanif10] **igmp ip-source-policy** 2001	（可选）在以上接口上配置设备根据源地址对 Report/Leave IGMP 报文进行过滤。参数{ *basic-acl-number* | **acl-name** *acl-name* }用来指定用于创建过滤策略的数字型 ACL 或命名型 ACL（**但仅支持基本 ACL**） 【说明】Report/Leave 报文封装在 IP 报文中，配置了本命令后，设备会检查封装了 IGMP Report/Leave 报文的 IP 报文头中的源地址。在定义 ACL 规则时，通过 **permit** 参数配置接口仅接收指定源地址范围的 Report/Leave 报文。如果 ACL 未定义规则，则接口缺省过滤掉所有源地址范围的 Report/Leave 报文。如果不配置 ACL 参数，IGMP Report/Leave 报文源地址的过滤规则如下。 （1）如果源地址和接收报文的接口地址在同一网段，或者源地址是 0.0.0.0，正常处理该报文 （2）如果源地址和接收报文的接口地址不在同一网段，则丢弃该报文 缺省情况下，交换机不对 Report/Leave 报文进行过滤，即处理所有接收到的 Report/Leave 报文，可用 **undo igmp ip-source-policy** 命令取消对 IGMP 报文源地址的过滤

13.1.4 配置 IGMP SSM Mapping

在SSM模型PIM-SM组播网络中，要求组播设备接口运行IGMPv3，但某些组播用户主机只能运行IGMPv1或IGMPv2。为了向这些用户同样提供SSM服务，需要在组播设备上配置SSM Mapping静态映射功能。

SSM Mapping是通过给 SSM组播组地址映射一个或多个组播源地址，将 IGMPv1或 IGMPv2 Report报文中（*，G）信息转换为一组（S，G）信息来实现SSM服务的。缺省情况下，SSM组地址范围为232.0.0.0～232.255.255.255，但可通过配置来扩展SSM组地址范围，具体将在本章后面介绍。配置SSM Mapping的具体步骤如表 13-4所示。

表13-4 SSM Mapping的配置步骤

步骤	命令	说明
1	**system-view** 例如：<HUAWEI> **system-view**	进入系统视图
2	**igmp** 例如：[HUAWEI] **igmp**	进入 IGMP 视图
3	**ssm-mapping** *group-address* { *group-mask* | *group-mask-length* } *source-address* 例如：[HUAWEI-igmp] **ssm-mapping** 224.0.5.5 24 10.10.10.1	配置静态 SSM 源组映射规则。命令中的参数说明如下。 （1）*group-address*：指定要映射的组播组 IP 地址，取值范围是 224.0.1.0～239.255.255.255 （2）*group-mask*：二选一参数，指定组播组 IP 地址的子网掩码 （3）*group-mask-length*：二选一参数，指定组播组 IP 地址的子网掩码长度 （4）*source-address*：指定要与以上组播组 IP 地址进行映射的组播源 IP 地址，是单播 IP 地址 缺省情况下，未配置 SSM 映射规则，可用 **undo ssm-mapping** { *group-address* { *mask* | *mask-length* } [*source-address*] | **static all** }命令删除指定的静态SSM源组映射规则，但尽量不要使用 **all** 选项，因为这样会将所有配置的 SSM 映射规则都清除
4	**quit** 例如：[HUAWEI-igmp] **quit**	退出 IGMP 视图，返回系统视图
5	**interface** *interface-type interface-number* 例如：[HUAWEI] **interface vlanif** 10	（可选）键入要配置 IGMP 查询器参数的 VLAN 或者 Loopback 接口（当然该接口必须已使能 IGMP），进入接口视图
6	**igmp ssm-mapping enable** 例如：[HUAWEI-Vlanif10] **igmp ssm-mapping enable**	（可选）在以上接口上使能 SSM Mapping 功能。只有在接口上使能 SSM Mapping，配置的 SSM 源/组地址映射表项才能生效 缺省情况下，接口未使能 SSM Mapping，可使用 **undo igmp ssm-mapping enable** 命令恢复缺省的去使能状态

13.1.5 配置 IGMP Limit

IGMP Limit提供了对组成员关系的个数限制功能。配置了组成员关系个数限制功能后，当收到IGMP报文时，首先判断是否超过配置的个数限制，如果没有超过就建立组成员关系，给用户转发该组的数据流。

组成员关系的计数规则如下。

（1）每个（*，G）组成员关系计为一个表项。

（2）每个（S，G）源组成员关系计为一个表项。

（3）使用 SSM Mapping的每个（*，G）组成员关系计为一个表项，按照映射生成的（S，G）表项不进行计数。

IGMP Limit功能可以在全局或者具体 IGMP接口上配置，具体配置步骤如表 13-5所示。

表13-5 IGMP Limit配置步骤

步骤	命令	说明
1	**system-view** 例如：<HUAWEI> **system-view**	进入系统视图
2	**igmp global limit** *number* 例如：[HUAWEI] **igmp global limit** 100	全局配置 IGMP 组成员关系个数限制，取值范围为 1～2 048 的整数（S7700/9300/9300E/9700 系列的取值范围为 1～49 512） 缺省情况下，整个交换机上可以创建的所有 IGMP 表项的最大个数为 2 048（S7700/9300/9300E/9700 系列的最大个数为 16 384），可用 **undo igmp global limit** 命令取消整个交换机上 IGMP 表项总和的最大个数限制 【说明】在 IGMP 视图下执行 **limit** *number* 命令也可配置全局 IGMP 组成员关系个数限制。如果同时配置，较小的取值生效
3	**interface** *interface-type interface-number* 例如：[HUAWEI] **interface vlanif** 10	（可选）键入要配置 IGMP Limit 功能的 VLAN 或者 Loopback 接口（当然该接口必须已使能 IGMP），进入接口视图
4	**igmp limit** *number* [**except** *acl-number*] 例如：[HUAWEI-Vlanif10] **igmp limit** 100 **except** 2001	（可选）配置当前接口上能够创建的组成员关系个数限制。命令中的参数说明如下。 （1）*number*：指定当前接口可以创建的 IGMP 表项最大值，取值范围为 1～2 048 的整数（S7700/9300/9300E/9700 系列的取值范围为 1～16 384） （2）**except** *acl-number*：可选参数，指定不受 *number* 参数限制的组播组范围，是通过 ACL 定义的。只对组地址进行过滤，则可使用基本 ACL，如果对（S，G）源组关系进行过滤，则要使用高级 ACL。如果没有使用本参数，则动态创建的所有组或源组时都受 IGMP 表项最大个数的限制；如果使用本参数，则先要配置相应的 ACL，接口将按照该 ACL 过滤收到的 IGMP Report 报文 缺省情况下，整个交换机上可以创建的所有 IGMP 表项的最大个数为 2 048，可用 **undo igmp limit** 命令删除当前接口可以维护 IGMP 组成员关系的最大个数限制

13.1.6 IGMP管理

配置好IGMP功能后，可以通过以下display任意视图命令查看相关IGMP信息，通过reset用户视图命令清除相关IGMP统计信息。

（1）使用display igmp interface [interface-type interface-number |up |down] [verbose] 命令查看指定接口或者状态为Up或者Down的IGMP接口上的详细信息（选择verbose可选项时）或摘要IGMP配置和运行信息。

（2）使用display igmp group [group-address | interface interface-type interface-number] * [verbose] 命令查看指定组播组或（和）指定 IGMP接口上动态加入的 IGMP组播组成员信息。

（3）使用以下命令查看静态IGMP组播组的成员信息。

① display igmpgroup [group-address] static [up |down] [verbose] 命令查看状态为Up或Down的IGMP接口的详细（选择verbose可选项时）或摘要信息。

② display igmpgroup [group-address] static interface-number 命令查看指定或所有IGMP静态组播组加入的接口数量。

③ display igmp group static interface interface-type interface-number entry-number命令查看指定IGMP接口加入的IGMP静态组播组或源组数量。

（4）执行display igmp group [group-address | interface interface-type interface-number] * [static] [verbose] 命令查看指定组播组或（和）指定 IGMP接口上的 IGMP组播组成员信息。

（5）使用display igmp control-message counters [interface interface-type interface- number] [message-type {query | report }] 命令查看指定或者所有 IGMP 接口上的IGMP报文统计计数。

（6）执行 display igmp routing-table [group-address [mask {group-mask |group-mask- length }] | source-address [mask { source-mask |source-mask-length }]] * [static] [outgoing- interface-number [number]] 命令查

看指定单个或一个范围的组播组的 IGMP路由表信息。但只有当接口上只运行了IGMP（没有运行PIM），且接口为IGMP查询器的情况下才会生成IGMP路由表项。

（7）使用display igmpgroup [group-address | interface interface-type interface-number] *ssm-mapping [verbose] 命令查看指定组播组或（和）指定 IGMP接口上配置的映射规则中的详细（选择verbose可选项时）或者摘要组播组信息。

（8）使用 display igmp ssm-mapping {group [group-address] | interface [interface-type interface-number] } 命令查看指定组播组或指定 IGMP 接口上配置的映射关系及 SSM Mapping运行状态。

（9）使用 reset igmpgroup {all | interface interface-type interface-number {all |group- address [mask { group-mask |group-mask-length }] [source-address [mask { source-mask |source-mask-length }]] } }用户视图命令清除指定IGMP接口或所有接口上加入的指定范围或者所有动态加入的组播组。

（10）使用undo igmp static-group {all | group-address [inc-step-mask { group-mask |group-mask-length }number group-number] [source source-address] }接口视图命令清除对应IGMP接口上指定范围或者所有静态加入的组播组。

13.1.7 IGMP基本功能配置示例

本示例拓扑结构如图13-1 所示，在主机侧存在两个主机网段N1 和 N2，HostA和 HostC 分别为 N1 和 N2 中的组播组成员。网络中传播组播数据使用的组播组地址为225.1.1.1～225.1.1.5，组播组成员HostA只购买了组225.1.1.1对应的节目，HostC则没有限制。

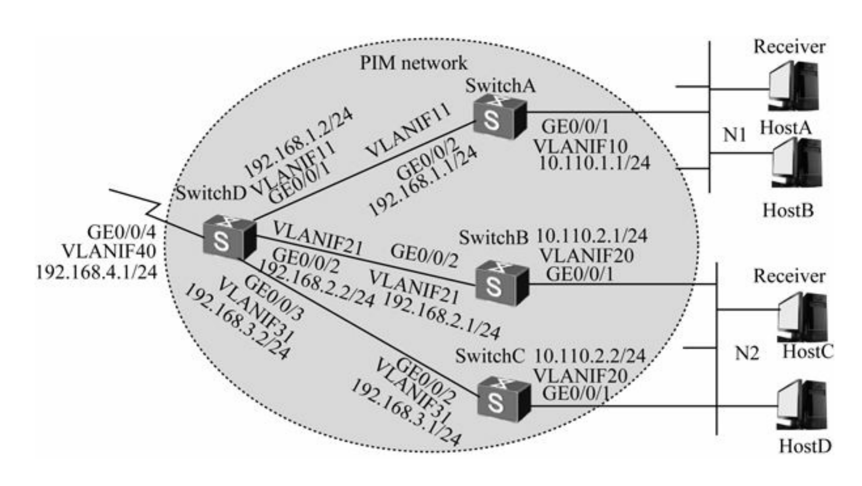


图13-1 IGMP的基本功能配置示例拓扑结构

1．基本配置思路分析

从图中的网络结构可以看出，本示例不仅涉及本章前面介绍的IGMP协议，还涉及本章后面要介绍的PIM协议，所以本示例同时涉及IGMP和PIM两种组播协议的基本PIM功能配置。下面是本示例的基本配置思路。

（1）配置网络中的单播路由协议（如可采用单播静态路由、RIP、OSPF之类的动态路由等），实现网络层互通。为了实现这一步，需要在各 Switch 的接口配置 IP 地址和单播路由协议。单播路由正常是组播路由协议正常工作的基础。本示例不作具体介绍。

（2）配置基本组播功能：全局使能组播路由功能，在与组播组成员连接的接口上使能PIM和IGMP，指定RP，以实现组播数据可以在网络中转发。

（3）通过 ACL 配置对 HostA 能接收的组播数据进行过滤，以实现示例中对 HostA接收的组播数据进行限制。

2．具体配置步骤

下面是本示例的具体配置步骤。

（1）配置各Switch接口IP地址和单播路由协议。

按照图13-1配置各VLAN接口的IP地址和掩码，并配置各Switch之间采用OSPF进行互连，确保网络中各Switch间能够在网络层互通。具体配置过程略。

（2）全局使能组播路由功能，在各组播交换机的所有接口上使能PIM-SM功能（需要先把物理接口加入到对应的VLAN中），并配置以SwitchD的VLANIF40为静态RP。因为SwitchA、SwitchB、SwitchC和SwitchD上的配置方法一样，所以下面仅以SwitchA为例进行介绍。

```
[SwitchA] vlan batch 10 11
[SwitchA] interface gigabitethernet 0/0/1
[SwitchA -GigabitEthernet0/0/1] port link-type access
[SwitchA -GigabitEthernet0/0/1] port default vlan 10
[SwitchA] interface gigabitethernet 0/0/2
[SwitchA -GigabitEthernet0/0/2] port link-type access
[SwitchA -GigabitEthernet0/0/2] port default vlan 12
[SwitchA -GigabitEthernet0/0/2] quit
[SwitchA] multicast routing-enable    #---全局使能组播路由功能
[SwitchA] interfacevlanif 10
[SwitchA-Vlanif10] pim sm    #---在VLAN10接口上（相当于在GE0/0/1接口上）启用PIM-SM
[SwitchA-Vlanif10] quit
[SwitchA] interfacevlanif 11
[SwitchA-Vlanif11] pim sm
[SwitchA-Vlanif11] quit
[SwitchA] pim
[SwitchA-pim] static-rp 192.168.4.1    #---配置SwitchD的GE0/0/4接口为静态RP
[SwitchA-pim] quit
```

（3）在SwitchA、SwitchB、SwitchC组播组成员侧接口上使能IGMP功能。现也仅以SwitchA的配置为例进行介绍，SwitchB和SwitchC上的配置过程与此类似，配置过程略。

```
[SwitchA] interface vlanif 10
[SwitchA-Vlanif10] igmp enable
[SwitchA-Vlanif10] quit
```

（4）通过 IGMP 报文过滤功能配置 SwitchA 的 VLANIF10 接口只能加入组播组225.1.1.1。要先创建一个允许以组播组地址225.1.1.1为源地址报文通过的基本ACL，然后在SwitchA的VLANIF10接口上应用该策略。

```
[SwitchA] acl number 2001
[SwitchA-acl-basic-2001] rule permit source 225.1.1.1 0
[SwitchA-acl-basic-2001] quit
[SwitchA] interface vlanif 10
```

[SwitchA-Vlanif10] igmp group-policy 2001

[SwitchA-Vlanif10] quit

配置好后，可以通过display igmp interface命令查看各接口上 IGMP的配置和运行情况，以验证配置结果。以下是SwitchA的VLANIF10接口上IGMP的显示信息。从中可以看出其基本配置，以及所应用的组策略。

```
<SwitchA>display igmp interface vlanif 10
Interface information
Vlanif 10(10.110.1.1):
    IGMP is enabled
    Current IGMP version is 2
    IGMP state: up
    IGMP group policy: 2001
    IGMP limit: -
    Value of query interval for IGMP (negotiated): -
    Value of query interval for IGMP (configured): 60 s
    Value of other querier timeout for IGMP: 0 s
    Value of maximum query response time for IGMP: 10 s
    Querier for IGMP: 10.110.1.1 (this router)
  Total 1 IGMP Group reported
```

13.1.8 静态加入组播组配置示例

本示例拓扑结构如图13-2所示，在主机侧存在两个主机网段N1和N2，N1中有一个组播组成员HostA，N2中有HostC和HostD两个组播组成员。现希望HostA长 期稳定地接收组播组 225.1.1.3 的数据，HostC 和 HostD 对所接收的组播组数据没有要求。

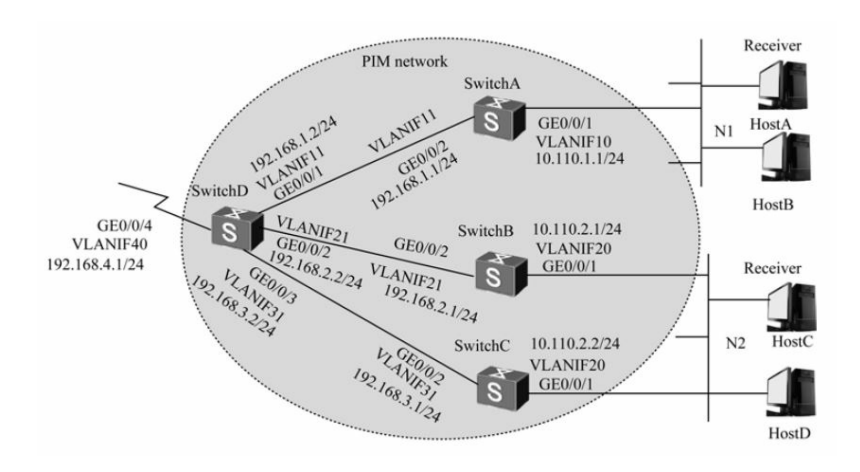


图13-2 静态加入组播配置示例拓扑结构

本示例与上节介绍的配置示例差不多，唯一不同的是上节示例介绍的是通过 IGMP报文过滤方式来限定HostA主机加入的组播组，而本示例介绍的是要求HostA接收主机静态加入组播组225.1.1.3。

正因如此，本示例的其他配置均可参见上节介绍，在此仅介绍HostA主机静态加入组播组225.1.1.3的配

置方法。具体如下。

[SwitchA] interfacevlanif 10

[SwitchA-Vlanif10] igmp static-group 225.1.1.3　#---静态加入到 IP地址为225.1.1.3的组播组中

[SwitchA-Vlanif10] quit

全部配置好后，可以通过使用display igmp group static命令查看接口上静态加入的组播组，以验证配置结果。

<SwitchA>display igmp group static

Static join group information

Total 1 entry, Total 1 active entry

Group Address	Source Address	Interface	State	Expires
225.1.1.3	0.0.0.0	Vlanif10	UP	never

13.1.9 IGMP SSM Mapping配置示例

本示例拓扑结构如图13-3所示，同时采用ASM和SSM模式提供组播服务。由于与组播组成员相连的Switch接口上运行IGMPv3，组播组成员主机上运行的是IGMPv2，且不能升级到IGMPv3，因此，该主机在加入组播组时无法指定组播源，必须依靠SSM Mapping来实现。

当前网络中的SSM组播组地址范围是232.1.1.0/24，Source1、Source2和Source3都向该范围内的组播组发送组播数据，而组播组成员只想接收来自 Source1 和 Source3 的组播数据。

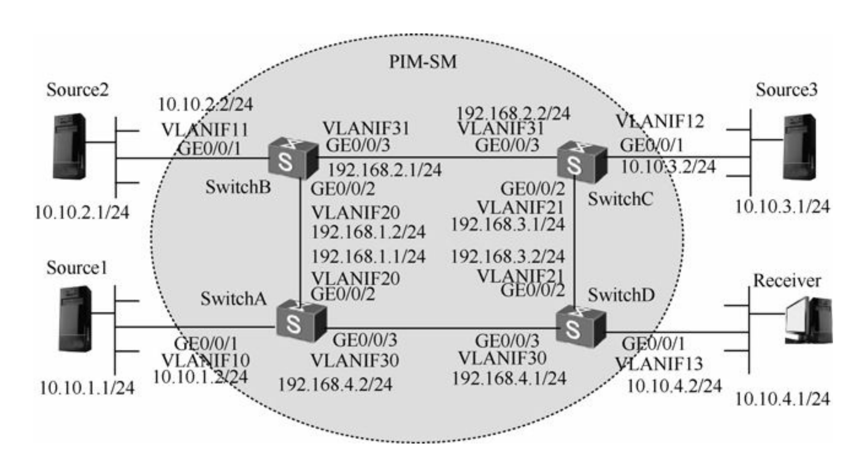


图13-3 SSM Mapping配置示例拓扑结构

1．基本配置思路分析

本示例有两项基本要求：一是通过SSM Mapping实现运行 IGMPv2的组播组成员可以使用SSM服务；二是通过SSM Mapping的组播组和组播源映射功能，使组播组成员仅可接收特定的指定源组播组数据。当然，首先要进行的也是一个基本组播网络的基本功能配置，如各交换机接口上的 PIM-SM 功能的使能，并指定相同的 RP，以及与组播组成员连接的交换机接口上的IGMP功能。

2．具体配置步骤

（1）配置IP地址和单播路由协议。按照图13-3标注配置各VLAN接口的IP地址和掩码，并配置各Switch之间采用OSPF进行互连，确保网络中各Switch间能够在网络层互通。具体配置过程略。

（2）在SwitchD上全局使能组播路由功能，并在各VLAN接口上配置PIM-SM，并在主机侧VLANIF13

接口上配置运行IGMPv3。

[SwitchD] multicast routing-enable

[SwitchD] interface vlanif 13

[SwitchD-Vlanif13] pim sm

[SwitchD-Vlanif13] igmp enable

[SwitchD-Vlanif13] igmp version 3

[SwitchD-Vlanif13] quit

[SwitchD] interface vlanif 21

[SwitchD-Vlanif21] pim sm

[SwitchD-Vlanif21] quit

[SwitchD] interfacevlanif 30

[SwitchD-Vlanif30] pim sm

[SwitchD-Vlanif30] quit

（3）在SwitchA、SwitchB和SwitchC上全局使能组播路由功能，并在各VLAN接口上使能 PIM-SM。因它们的配置基本一样，所以下面仅以 SwitchA 为例进行介绍， SwitchB和SwitchC的配置方法参见即可，配置过程略。

[SwitchA] multicast routing-enable

[SwitchA] interfacevlanif 10

[SwitchA-Vlanif10] pim sm

[SwitchA-Vlanif10] quit

[SwitchA] interface vlanif 20

[SwitchA-Vlanif20] pim sm

[SwitchA-Vlanif20] quit

[SwitchA] interfacevlanif 30

[SwitchA-Vlanif30] pim sm

[SwitchA-Vlanif30] quit

（4）在SwitchD上配置VLANIF30为C-BSR和C-RP。因为本网络中只配置了一个C-BSR和一个C-RP，所以VLANIF30最终会直接成为BSR和RP。

[SwitchD] pim

[SwitchD-pim] c-bsrvlanif 30

[SwitchD-pim] c-rp vlanif 30

[SwitchD-pim] quit

（5）在SwitchD的VLANIF13上使能SSM Mapping功能。

[SwitchD] interfacevlanif 13

[SwitchD-Vlanif13] igmp ssm-mapping enable

[SwitchD-Vlanif13] quit

（6）在所有 Switch 上配置 SSM 组播组地址范围，以限定组播数据的发送。因为SwitchA、SwitchB、SwitchC和SwitchD上的配置方法一样，所以下面仅SwitchA的配置为例进行介绍。

[SwitchA] acl number 2000

[SwitchA-acl-basic-2000] rule permit source232.1.1.0 0.0.0.255

```
[SwitchA-acl-basic-2000] quit
[SwitchA] pim
[SwitchA-pim] ssm-policy 2000
[SwitchA-pim] quit
```

（7）在连接主机的Switch上配置SSM Mapping映射规则，将 232.1.1.0/24范围内的组播组映射到组播源Source1和Source3上，以实现组播组成员接收到Source1和Source3发来的组播数据。

```
[SwitchD] igmp
[SwitchD-igmp] ssm-mapping 232.1.1.0 24 10.10.1.1
[SwitchD-igmp] ssm-mapping 232.1.1.0 24 10.10.3.1
[SwitchD-igmp] quit
```

配置好后，可通过display igmp ssm-mapping group命令查看Switch上源和组的映射关系，以验证配置结果。

```
<SwitchD>display igmp ssm-mapping group
IGMP SSM-Mapping conversion table
Total 2 entries        2 entries matched
00001. (10.10.1.1, 232.1.1.0/24)
00002. (10.10.3.1, 232.1.1.0/24)
Total 2 entries matched
```

还使用 display igmp group ssm-mapping命令查看 Switch特定源/组地址的信息。SwitchD上特定源/组地址信息显示如下，从中可以看出组播组成员已加入到组232.1.1.1中。

```
<SwitchD>display igmp group ssm-mapping
IGMP SSM mapping interface group report information
Limited entry of this VPN-Instance: -
Vlanif13 (10.10.4.2):
Total 1 IGMP SSM-Mapping Group reported
Group Address      Last Reporter      Uptime      Expires
232.1.1.1          10.10.4.1          00:01:44    00:00:26
```

13.1.10 IGMP Limit配置示例

本示例拓扑结构如图13-2所示，该网络中的用户主机通过组播方式接收视频节目。现假设与SwitchA相连网段的HostA上订购了一个长期的组地址为225.1.1.3的节目，要求当网络中的用户主机点播的节目数量达到限制值时不允许再点播新的节目，保证用户已订购节目的接收质量。

1．基本配置思路分析

本示例有两项主要配置。

（1）为组播组成员HostA配置静态加入组播组225.1.1.3，使该用户能长期接收发往组播组225.1.1.3的数据。

（2）采用 IGMP Limit功能来限制连接订购节目的用户的交换机上配置的组成员关系数量，以保证用户已订购节目的接收质量。

其他的组播网络基本功能配置与13.1.7节介绍的示例差不多。

2．具体配置步骤

（1）配置各Switch接口IP地址和单播路由协议。按照图13-4中的标注，配置各VLAN接口的IP地址和掩码，并配置各Switch之间采用OSPF进行互连，确保网络中各Switch间能够在网络层互通。具体配置过程略。

（2）全局使能组播路由功能，并在所有 VLAN 接口上使能 PIM-SM 功能，同时以SwitchD上的GE0/0/4接口为静态RP。因为SwitchA、SwitchB、SwitchC和SwitchD上的配置方法都一样，所以下面仅以SwitchA为例进行介绍。

[SwitchA] multicast routing-enable

[SwitchA] interfacevlanif 10

[SwitchA-Vlanif10] pim sm

[SwitchA-Vlanif10] quit

[SwitchA] interfacevlanif 11

[SwitchA-Vlanif11] pim sm

[SwitchA-Vlanif11] quit

[SwitchA] pim

[SwitchA-pim] static-rp 192.168.4.1

[SwitchA-pim] quit

（3）配置SwitchA、SwitchB和SwitchC的组播组成员侧接口使能IGMP，同样因为它们的配置方法一样，下面也仅以SwitchA为例进行介绍。

[SwitchA] interfacevlanif 10

[SwitchA-Vlanif10] igmp enable

（4）将SwitchA的组播组成员侧接口静态加入组播组225.1.1.3，使用户能长期接收发往组播组225.1.1.3的数据。

[SwitchA-Vlanif10] igmp static-group 225.1.1.3

[SwitchA-Vlanif10] quit

（5）在连接已订购节目用户的最后一跳交换机上配置IGMP组成员关系个数限制。本示例是需要在SwitchA上配置，假设总共可以创建50个IGMP组成员关系。

[SwitchA] igmp global limit 50

还可在具体的接口（如VLANIF10）上配置总共可以创建的IGMP组成员关系数量，假设为30个（肯定要小于SwitchA上全局的成员关系限制数）。

[SwitchA] interfacevlanif 10

[SwitchA-Vlanif10] igmp limit 30

[SwitchA-Vlanif10] quit

如果还要保证SwitchB和SwitchC上已订购节目用户的接收质量，可以按照上面介绍的SwitchA上HostA配置方法配置静态加入组播组，并且配置 IGMP Limit功能。

配置好后，可通过使用display igmp interface命令查看交换机接口上 IGMP的配置和运行情况。SwitchA的VLANIF10接口上IGMP的显示信息如下，从中可以看到SwitchA的VLANIF10上可创建的IGMP组成员关系的最大个数为30个。

<SwitchA>display igmp interface vlanif 10

Interface information

vlanif10(10.110.1.1):

IGMP is enabled
Current IGMP version is 2
IGMP state: up
IGMP group policy: none
IGMP limit: 30
Value of query interval for IGMP (negotiated): -
Value of query interval for IGMP (configured): 60 s
Value of other querier timeout for IGMP: 0 s
Value of maximum query response time for IGMP: 10 s
Querier for IGMP: 10.110.1.1 (this router)

13.2 PIM-DM（IPv4）配置与管理

在 PIM-DM 模式中使用“推”（Push）模式转发组播报文，就是由 PIM 路由器向组播成员主动推送组播数据，因为 PIM-DM 网络仅适用于 ASM 模型，即它的组播源是任意的，组播成员和组播路由器都不关心组播源的位置。当网络中有活跃的组播源出现，即有组播源需要向某组播组发送组播数据时，会将组播数据扩散到全网，借助RPF 检查机制创建组播路由表项，实现组播数据转发。因为组播数据要在全网泛洪扩散，所以一般用于规模较小、组成员分布密集的组播网络，否则可能造成组播路由器的数据转发压力。

PIM-DM的关键工作机制包括邻居发现、扩散、剪枝、嫁接、断言和状态刷新这几个过程。如图13-4所示，在PIM-DM网络中出现了活跃的组播源Source后，通过组成员管理协议IGMP，SwitchD、SwitchE了解到与其相连的HostB、HostC为组播组成员，便将接收到的组播数据分别向组成员所在网段转发。由于与SwitchC相连网段没有组成员，它会逐跳向上游发起剪枝操作。组播源的最后一跳PIM组播路由器SwitchA接收到剪枝报文后，就会将与SwitchC相连的下游接口从PIM路由表项的下游接口列表中删除，抑制组播数据向该接口的网段转发。PIM-DM 网络就是这样通过这种周期性的“扩散-剪枝”，来构建并维护一棵单向无环的SPT（Source Specific Shortest Path Tree，源指定最短路径树）。整个 **SPT** 路径是以组播源为起点的 ，同时组播源也是组播的中心。具体 **PIM-DM**工作原理参见本书第**12**章**12.3.2**节 。

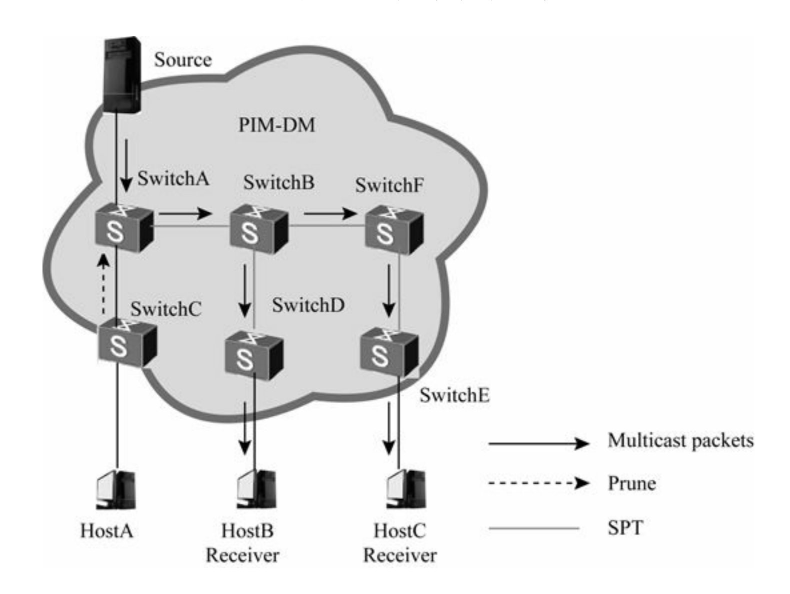


图13-4 PIM-DM网络构建SPT的示例

13.2.1 PIM-DM（IPv4）特性的产品支持

除**S2700/3700**系列外，其他所有S系列交换机均支持PIM-DM。

PIM-DM可配置的主要特性包括基本功能的配置、PIM-DM控制参数的调整和PIM Silent。但除了基本功能配置外，其他均为可选配置任务 。

1．PIM-DM基本功能

PIM-DM基本功能就是对对应的VLAN或者Loopback接口上使能PIM-DM功能。在接口上使能了PIM-DM功能之后，系统采用PIM-DM的缺省值就可以正常工作，将组播源发出的组播数据分发到组成员网段。

2．PIM-DM控制参数

有时候为了提高网络安全性，增强组播报文转发的控制能力，根据实际需要，设备支持调整如表13-6所示的PIM-DM控制参数。

表13-6 PIM-DM控制参数

参数	说明
组播源控制参数	设备可以基于组播源来控制组播报文的转发。一方面有助于数据流量控制，另一方面可以限定下游组播组成员能够获得的信息，提高安全性
邻居控制参数	设备间通过交互 Hello 报文建立 PIM 邻居关系，协商各类控制参数，所以可以基于 Hello 报文来控制邻居间的关系。同时可以配置灵活的邻居控制策略，提高网络安全性，比如防止非法 PIM 邻居的入侵
剪枝控制参数	如果当前与设备相连的网段没有组播组成员，设备就需要向上游发送剪枝报文，请求停止转发组播数据。可根据实际需要调整剪枝过程的控制参数，控制组播报文转发来支持不同转发场景
嫁接控制参数	为使被剪枝网段中出现的组播组成员在下一次“扩散-剪枝”来临前就接收到组播数据，设备需要向上游发送嫁接报文请求恢复转发组播数据。可根据实际需要调整嫁接过程的控制参数，控制组播报文转发来支持不同转发场景

（续表）

参数	说明
状态刷新控制参数	周期性的“扩散-剪枝”将造成很大的网络资源浪费。为防止被剪枝接口因为剪枝状态超时而恢复转发，系统启用了状态刷新功能，周期性地发送 State-Refresh（状态刷新）报文，刷新接口剪枝定时器，使没有组播数据转发需求的接口一直处于抑制状态
断言控制参数	当设备从下游接口接收到组播数据时，说明该网段中还存在其他的上游设备。设备从该接口发出断言报文，参与竞选唯一上游的转发者。可根据实际需要调整发送断言的间隔，来控制断言竞选的周期

3．PIM Silent

在直连用户主机的接口上需要使能PIM协议，以便在该接口上可以建立PIM邻居，处理各类 PIM 协议报文。但这里存在一个问题，那就是如果有恶意主机模拟向 PIM 路由器发送PIM Hello报文时，有可能导致设备瘫痪。为了避免这样的情况发生，可以将该接口设置为 PIM Silent状态（即 PIM消极状态），这样就可禁止该接口接收和转发任何PIM协议报文，删除该接口上的所有PIM邻居以及PIM状态机。但该接口上的IGMP功能不受影响。

该功能仅适用于与用户主机网段直连的 PIM 设备接口，且该用户网段只与这一台**PIM**设备相连 。配置了该功能后，接口将不再接收和转发任何PIM协议报文，但这样配置后，该接口配置的其他 PIM 功能将失效，需谨慎使用。如果用户网段与多台 PIM 设备相连，且如果在多个PIM设备接口上配置PIM Silent，则这些接口都成为了静态DR，会导致该网段中同时存在多个DR，从而引发组播故障。

S2700/3700系列交换机不支持该特性 。

与 PIM-DM 相关的功能和参数缺省配置为 PIM-DM 未使能、状态刷新功能在使能PIM-DM功能后使能，PIM Silent功能未使能。

13.2.2 配置PIM-DM基本功能

在配置PIM-DM前需要先配置好单播路由协议，保证网络内单播路由畅通。且设备上不能同时使能PIM-DM和PIM-SM，如果接口上需要同时使能**PIM-DM**和**IGMP**，必须先使能**PIM-DM**，再使能**IGMP** 。

PIM-DM基本功能的配置很简单，主要就是两项基本任务：一是全局使能组播路由功能（如果在配置其他组播功能时已使能，则无需再使能），二是在对应的 VLAN 接口或者Loopback接口上使能PIM-DM功能。具体配置步骤如表13-7所示。

表13-7 PIM-DM基本功能配置步骤

步骤	命令	说明
1	**system-view** 例如：<HUAWEI> **system-view**	进入系统视图
2	**multicast routing-enable** 例如：[HUAWEI] **multicast routing-enable**	全局使能组播路由功能。其他说明参见 13.1.2 节表 13-1 中的第 2 步
3	**interface** *interface-type interface-number* 例如：[HUAWEI] **interface vlanif** 10	键入要配置 PIM-DM 功能的 VLAN 或者 Loopback 接口，进入接口视图
4	**pim dm** 例如：[HUAWEI-Vlanif10] **pim dm**	在以上接口上使能 PIM-DM 功能。在接口上使能了 PIM-DM 功能后，交换机才能与相邻的设备建立 PIM 邻居，对来自 PIM 邻居的协议报文进行处理 缺省情况下，接口上未使能 PIM-DM，可使用 **undo pim dm** 命令恢复缺省的去使能状态

（续表）

13.2.3 调整组播源控制参数

每当PIM设备在接收到源S发往组播组G的组播报文后，就会启动该（S，G）表项的定时器，即源生存时间（缺省值为 **210s** ）。下次如果超时前接收到该组播源发来的报文，则重置定时器；如果超时后没有接收到该组播源发来的报文，则认为该（S，G）表项失效，将其删除。通过这种方法可以及时地更新PIM路由器上的组播转发表项。

另外，如果希望控制组播流量或者保证组播成员所接收的组播数据的安全性，还可在PIM设备上配置源地址过滤策略，只接收该策略允许范围内组播源发送的组播数据，拒绝非法的组播数据。缺省情况下，没有过滤策略，即接收任何组播源发来的组播数据。通过基本或高级 ACL 可对组播源地址或组址进行过滤，还可对组播源生存时间进行控制，提高数据安全性、控制网络流量。但在调整组播源控制参数之前，需配置好PIM-DM基本功能。具体配置步骤如表13-8所示。

表13-8 调整组播源控制参数的配置步骤

步骤	命令	说明
1	**system-view** 例如：<HUAWEI> **system-view**	进入系统视图
2	**pim** 例如：[HUAWEI] **pim**	进入 PIM 视图。可用 **undo pim** 命令清除 PIM 视图下进行的配置，将删除所有 IPv4 PIM 全局配置信息，请慎用
3	**source-lifetime** *interval* 例如：[HUAWEI-pim] **source-lifetime** 120	配置组播源生存时间（**适用于所有组播源**），取值范围为 60～65 535 的整数秒 接口第一次收到源 S 发出的组播报文后，启动定时器；然后，每接收到 S 发出的组播报文就重置定时器；如果定时器超时，则认为（S，G）表项失效 缺省情况下，超时时间是 210s，可用 **undo source-lifetime** 命令恢复时间间隔为缺省值
4	**source-policy** { *acl-number* | **acl-name** *acl-name* } 例如：[HUAWEI-pim] **source-policy** 2001	配置源地址过滤策略，使交换机对接收的组播数据报文根据指定数字型（选择 *acl-number* 参数时）或者命名型（选择 *acl-name* 参数时）ACL（可以是基本 ACL，也可以是高级 ACL）所限定的源或源组（即组播源和组播组）进行过滤，防止非法源信息传播到 PIM 网络。但通过本命令配置源过滤策略可限定合法的组播源或者组播源组地址范围，**所有未通过该过滤规则的报文将被丢弃，但不过滤静态加入的（S，G）** 如果配置的是基本 ACL，规则中的源地址代表的是组播源地址，即只转发组播源地址属于源地址过滤规则允许范围的组播报文；如果配置的是高级 ACL，规则中源地址代表组播源地址，目的地址为组播组地址，即只转发组播源地址和组播组地址，都属于过滤规则允许范围内的组播报文；如果指定 ACL 没有配置过滤规则，则不转发任何组播源/组地址发送的组播报文 缺省情况下，交换机不根据组播源或组播源组地址过滤组播数据报文，可用 **undo source-policy** 命令删除过滤配置

【示例 1】使用数字型ACL配置接收组播源地址为10.10.1.2的组播数据包，丢弃组播源地址为10.10.1.1的组播数据包。

```
<HUAWEI> system-view
[HUAWEI] acl number2001
[HUAWEI-acl-basic-2001] rule permit source 10.10.1.2 0
[HUAWEI-acl-basic-2001] rule deny source10.10.1.1 0
[HUAWEI-acl-basic-2001] quit
[HUAWEI] multicast routing-enable
[HUAWEI] pim
[HUAWEI-pim] source-policy 2001
```

【示例 2】使用命名型ACL配置接收组播源地址为10.10.1.2的组播数据包，丢弃组播源地址为10.10.1.1的组播数据包。

```
<HUAWEI> system-view
[HUAWEI] acl name myacl
[HUAWEI-acl-adv-myacl] rule permit ip source 10.10.1.2 0
[HUAWEI-acl-adv-myacl] rule deny ip source10.10.1.1 0
[HUAWEI-acl-adv-myacl] quit
[HUAWEI] multicast routing-enable
[HUAWEI] pim
[HUAWEI-pim] source-policy acl-namemyacl
```

【示例 3】使用数字型ACL配置接收组播源地址为10.10.1.2，组播组地址为225.1.1.3的组播数据包。

```
<HUAWEI> system-view
[HUAWEI] acl number3001
[HUAWEI-acl-basic-3001] rule permit source 10.10.1.2 0destination 225.1.1.3 0
[HUAWEI-acl-basic-3001] quit
```

[HUAWEI] multicast routing-enable

[HUAWEI] pim

[HUAWEI-pim] source-policy 3001

13.2.4 调整邻居控制参数

通过调整邻居控制参数，控制邻居间 Hello 报文的交互，可以防止非法邻居关系的建立，保证PIM-DM网络的安全。在调整邻居控制参数之前，也需要先完成PIM-DM基本功能配置。

1．调整Hello报文的时间控制参数

PIM设备通过周期性地发送Hello报文来维护PIM邻居关系，就像RIP、OSPF这些动态路由协议中的邻居维护一样。当PIM设备收到邻居发来Hello报文后会启动定时器，时间设为该Hello报文的保持时间。如果超时后没有收到邻居发来的Hello报文则认为该邻居失效或者不可达。为了避免多个PIM设备同时发送Hello报文而导致冲突，当PIM设备接收到Hello报文时将延迟一段时间再发送Hello报文。该段时间的值为一个随机值，并且小于触发Hello报文的最大延迟。

发送Hello报文的时间间隔（缺省为**30s** ）、Hello报文的保持时间（缺省为**105s** ）在全局PIM视图下和接口视图下都可配置。如果同时配置，接口视图上的配置生效。但触发**Hello**报文的最大延迟时间 （缺省为**5s** ）只能在接口上配置 。

2．配置邻居过滤策略

设备支持以下两种邻居过滤策略，来保证PIM-DM网络的安全和畅通。

（1）限定合法的邻居地址范围，防止非法邻居入侵等。

（2）拒绝接收无Generation ID的Hello报文，保证与设备相连的都是正常工作的PIM邻居。

以上两项基本配置任务的具体配置步骤如表13-9所示，但Hello报文的时间控制参数、邻居过滤策略配置时并无先后顺序，可根据实际需要进行调整。

表13-9 调整邻居控制参数的配置步骤

步骤	命令	说明	
1	**system-view** 例如：<HUAWEI> **system-view**	进入系统视图	
2	**pim** 例如：[HUAWEI] **pim**	进入 PIM 视图。可用 **undo pim** 命令清除 PIM 视图下进行的配置，将删除所有 IPv4 PIM 全局配置信息，请慎用	全局调整邻居控制对数
3	**timer hello** *interval* 例如：[HUAWEI-pim] **timer hello** 100	全局配置发送 Hello 报文的时间间隔，取值范围为 1～18 000 的整数秒 缺省情况下，交换机发送 Hello 报文的时间间隔是 30s，可用 **undo timer hello** 命令恢复时间间隔为缺省值	
4	**hello-option holdtime** *interval* 例如：[HUAWEI-pim] **hello-option holdtime** 200	全局配置 Hello 报文的保持时间，即配置交换机等待接收其 PIM 邻居发送 Hello 报文的超时时间，取值范围为 1～65 535 的整数秒。如果超时后，设备没有收到该邻居后续发来的 Hello 报文，则认为邻居失效或不可达 Hello 报文的保持时间应该大于上一步配置的 Hello 报文发送间隔 缺省情况下，交换机等待接收其 PIM 邻居发送 Hello 报文的超时时间是 105s，可用 **undo hello-option holdtime** 命令恢复配置参数的缺省值	
5	**quit** 例如：[HUAWEI-pim] **quit**	退出发 PIM 视图，返回系统视图	在接口上调整邻居控制参数
6	**interface** *interface-type interface-number* 例如：[HUAWEI] **interface vlanif 100**	（可选）键入要配置邻居控制参数的 PIM-DM VLAN 接口或者 Loopback 接口，进入接口视图	
7	**pim timer hello** *interval* 例如：[HUAWEI-Vlanif100] **pim timer hello** 100	（可选）在以上接口上配置发送 Hello 报文的时间间隔，取值范围为 1～18 000 的整数秒 缺省情况下，交换机发送 Hello 报文的时间间隔是 30s，可用 **undo pim timer hello** 命令恢复时间间隔为缺省值	
8	**pim hello-option holdtime** *interval* 例如：[HUAWEI-Vlanif100] **pim hello-option holdtime** 300	（可选）在以上接口上配置 Hello 报文的保持时间，取值范围为 1～65 535 的整数秒 Hello 报文的保持时间应该大于上一步配置的 Hello 报文发送间隔 缺省情况下，交换机等待接收其 PIM 邻居发送 Hello 报文的超时时间是 105s，可用 **undo pim hello-option holdtime** 命令恢复配置参数的缺省值	
9	**pim triggered-hello-delay** *interval* 例如：[HUAWEI-Vlanif100] **pim triggered-hello-delay** 3	（可选）在以上接口上配置触发 Hello 报文的最大延迟，取值范围为 1～5 的整数秒。为了避免多个 PIM 设备同时发送 Hello 报文而导致冲突，当 PIM 路由器检测到网络中已存在 Hello 报文时，将自动选取小于本命令配置值的任意随机数进行延时，然后发送 Hello 报文 缺省情况下，触发 Hello 报文的最大延迟是 5s，可使用 **undo pim triggered-hello-delay** 命令恢复触发 Hello 报文的最大延迟为缺省值	

（续表）

步骤	命令	说明	
10	**pim neighbor-policy** *basic-acl-number* 例如：[HUAWEI-Vlanif100]**pim neighbor-policy** 2010	（可选）过滤以上接口上的 PIM 邻居。参数用来定义限制 PIM 邻居（单播 IP 地址）的基本 ACL，取值范围为 2 000～2 999。在定义 ACL 规则时，通过 permit 选项配置接口仅接收指定地址范围的 Hello 报文。如果 ACL 未定义规则，则接口过滤掉所有地址范围的 Hello 报文 【说明】为了防止某些非法邻居参与 PIM 协议，可通过执行此命令配置邻居过滤规则，限定合法的邻居地址范围，只与符合过滤规则的邻居建立邻居关系，删除不符合过滤规则的邻居。设备上配置了合法的邻居地址范围后，如果之前与其建立好邻居关系的 PIM 设备不在其合法地址范围内，后续将不会再收到邻居设备的 Hello 报文。邻居关系也会因 Hello 报文的保持时间超时而解除 缺省情况下，不过滤接口上的 PIM 邻居，可用 **undo pim neighbor-policy** 命令恢复缺省配置	配置邻居过滤策略
11	**pim require-genid** 例如：[HUAWEI-Vlanif100]**pim require-genid**	（可选）配置以上 PIM 接口拒绝无 Generation ID 参数的 Hello 报文 【说明】正常情况下，在接口上使能 PIM 后，设备会生成一个随机数作为 Hello 报文的 Generation ID。如果设备的状态有变化则生成新的 Generation ID。当对端设备接收到该 Hello 报文后，发现其中包含的 Generation ID 已改变，则认为 PIM 邻居的状态已经改变。执行此命令可配置设备拒绝接收无 Generation ID 的 Hello 报文，保证连接的 PIM 邻居都处于正常工作状态 缺省情况下，PIM 接口接收无 Generation ID 参数的 Hello 报文，可用 **undo pim require-genid** 命令恢复缺省配置	
12	**quit** 例如：[HUAWEI-Vlanif100] **quit**	退出接口视图，返回系统视图	
13	**pim** 例如：[HUAWEI] **pim**	进入 PIM 视图	
14	**neighbor-check** { **receive** | **send** } 例如：[HUAWEI] **neighbor-check receive**	使能 PIM 邻居发送（选择 **send** 二选一选项）或接收（选择 **receive** 二选一选项）的检查功能。加入和剪枝动作都是通过设备逐跳向上游 PIM 邻居发送 Join-Prune 报文完成；而断言竞选也是发生在多个 PIM 邻居之间。因此，有时候为了减少设备资源浪费，以及保证上述两类协议报文的安全性，可通过执行此命令配置邻居检查功能 设备上可同时配置设备的邻居发送和接收的检查功能。缺省情况下，没有使能 PIM 邻居检查功能，可用 **undo neighbor-check** 命令恢复缺省配置	

【示例 】配置VLANIF100与地址为4.4.4.4的交换机建立PIM邻居。

```
<HUAWEI>system-view
[HUAWEI] acl number2001
[HUAWEI-acl-basic-2001] rule permit source 4.4.4.4 0.0.0.0
[HUAWEI-acl-basic-2001] quit
[HUAWEI] multicast routing-enable
[HUAWEI] interface vlanif 100
[HUAWEI-Vlanif100] pim neighbor-policy 2001
```

13.2.5 调整剪枝控制参数

如果当前与设备相连的网段没有组播组成员，设备就需要向上游发送剪枝报文，请求停止转发组播数据。可根据实际需要调整剪枝过程的控制参数，控制组播报文转发来支持不同转发场景。但如果没有特殊需要，推荐使用缺省值。同样，在调整剪枝控制参数之前，需要完成PIM-DM基本功能配置任务。

在剪枝控制参数调整过程中，可以配置：Join-Prune报文的时间控制参数、Join-Prune报文的信息携带能力、剪枝延迟时间这3个方面，但它们的配置无先后顺序，也不是必须全部配置，用户可根据实际需要进行调整。

1．调整Join-Prune报文的时间控制参数

PIM 设备通过向上游发送的剪枝信息被封装在 PIM 协议通用的转发控制报文（即Join-Prune 报文）中。上游设备在收到 Join-Prune 报文后，就会启动定时器（缺省值为**210s** ），时间设为Join-Prune报文自身携带的保持时间。超时后，如果没有收到下游后续发来的Join-Prune报文，则恢复相应组播组下游接口的转发。

Join-Prune报文的保持时间在全局PIM视图下和接口视图下都可配置。如果同时配置，接口视图上的配置生效。

2．调整Join-Prune报文的信息携带能力

在PIM-DM网络中，Join-Prune报文主要包含了需要剪枝的表项信息。设备支持通过配置Join-Prune报文长度、包含表项数目、发送方式，来调整向上游发送剪枝信息的信息量。

（1）当PIM邻居设备性能比较差，处理单个Join-Prune报文耗时比较长，可以通过调整发送的 Join-Prune报文长度（缺省值为 **8 100**字节 ）来控制发送 Join-Prune报文携带的（S，G）表项数量，来降低PIM邻居设备的压力。

（2）当 PIM 邻居设备端口带宽较小时，可以通过调整周期性报文发送队列长度，控制每次发给PIM邻居设备的（S，G）表项数量（缺省值为**1 020**个 ），采取小量多批次方式发送Join-Prune报文，从而避免PIM邻居设备来不及处理就将报文丢弃，引起路由振荡。

（3）缺省情况下，为了提高发送效率，Join-Prune 报文都是打包向上游发送。如果不希望Join-Prune报文打包发送，可去使能此功能，使Join-Prune报文一个个地发送。

3．调整剪枝延迟时间

在PIM剪枝过程中，从收到下游设备发来的剪枝信息到继续向上游设备发送剪枝信息会有一段延迟时间，这个时间称为LAN-Delay（缺省值为**500ms** ）。PIM设备在向上游发完剪枝信息后，也不会立即将相应下游接口剪掉，还会保持一段时间向下游转发，以便下游设备有时间提出否决剪枝的请求。这段否决剪枝的时间称为Override-Interval（缺省值为 **2 500ms** ）。所以，实际上 PIM设备从收到剪枝信息到完成剪枝动作总共延迟了LAN-Delay＋Override-Interval = PPT。PPT表示当前交换机从收到下游剪枝报文到执行剪枝操作（抑制下游接口转发）之间的延时。在PPT时间内如果收到下游发来的剪枝否决报文，则取消剪枝操作。

LAN-Delay、Override-Interval在全局PIM视图下和接口视图下都可配置。如果同时配置，接口视图下的配置优先级高于系统视图下的配置，接口视图下的配置生效。

以上三方面配置任务的具体配置步骤如表13-10所示。

表13-10 调整剪枝控制参数的配置步骤

配置任务	步骤	命令	说明	
公共配置	1	**system-view** 例如：<HUAWEI> **system-view**	进入系统视图	
调整 Join-Prune 报文的时间控制参数	2	**pim** 例如：[HUAWEI] **pim**	进入 PIM 视图	全局配置
	3	**timer join-prune** *interval* 例如：[HUAWEI-pim] **timer join-prune** 80	配置向上游设备周期性发送 Join-Prune 报文的时间间隔，取值范围是 1～18 000 的整数秒 【说明】PIM 交换机通过向上游发送加入信息请求转发组播数据，发送剪枝信息请求停止转发组播数据。实际上，加入信息和剪枝信息都被封装在了 Join-Prune 报文中，PIM 路由器会周期性地将 Join-Prune 报文发送给上游设备来更新转发状态。可通过此命令设置 Join-Prune 报文的发送周期 该命令配置的时间间隔必须小于下一步 **holdtime join-prune** 命令配置的时间间隔，即发送 Join-Prune 报文的周期必须小于 Join-Prune 报文的保持时间 缺省情况下，向上游设备周期性发送 Join-Prune 报文的时间间隔是 60s，可用 **undo timer join-prune** 命令恢复全局发送 Join-Prune 报文的时间间隔为缺省值	
	4	**holdtime join-prune** *interval* 例如：[HUAWEI-pim] **holdtime join-prune** 1000	全局配置 Join-Prune 报文的保持时间，取值范围为 1～65 535 的整数秒。接收到 Join-Prune 报文的交换机依据该报文自身携带的保持时间来确定对应下游接口保持加入或剪枝状态的时间（由此可以看出，这一个参数值定义了两个相同的时间） 缺省情况下，Join-Prune 报文的保持时间是 210s，可用 **undo holdtime join-prune** 命令恢复全局的 Join-Prune 报文保持时间为缺省值	
	5	**quit** 例如：[HUAWEI-pim] **quit**	退出发 PIM 视图，返回系统视图	在接口上配置
	6	**interface** *interface-type interface-number* 例如：[HUAWEI] **interface vlanif 100**	（可选）键入要调整 Join-Prune 报文的时间控制参数的 PIM-DM VLAN 接口或者 Loopback 接口，进入接口视图	
	7	**pim timer join-prune** *interval* 例如：[HUAWEI-Vlanif100] **pim timer join-prune** 100	（可选）在接口上配置向上游设备周期性发送 Join-Prune 报文的时间间隔，取值范围为 1～18 000 的整数秒 该命令配置的时间间隔必须小于下一步的 **pim holdtime join-prune** 命令配置的时间间隔，即发送 Join-Prune 报文的周期必须小于 Join-Prune 报文的保持时间 缺省情况下，接口向上游设备周期性发送 Join-Prune 报文的时间间隔是 60s，可用 **undo pim timer join-prune** 命令恢复接口上发送 Join-Prune 报文的时间间隔为缺省值	

（续表）

配置任务	步骤	命令	说明	
调整 Join-Prune 报文的时间控制参数	8	**pim holdtime join-prune** *interval* 例如：[HUAWEI-Vlanif100]**pim holdtime join-prune** 500	（可选）在接口上配置 Join-Prune 报文的保持时间，取值范围为 1～65 535 的整数秒 缺省情况下，Join-Prune 报文的保持时间是 210s，可用 **undo pim holdtime join-prune** 命令恢复接口的 Join-Prune 报文保持时间为缺省值	在接口上配置
调整 Join-Prune 报文的信息携带能力	9	**quit** 例如：[HUAWEI-Vlanif100] **quit**	退出接口视图，返回系统视图	
	10	**pim** 例如：[HUAWEI] **pim**	进入 PIM 视图	
	11	**join-prune max-packet-length** *packet-length* 例如：[HUAWEI-pim]**join-prune max-packet-length** 1500	配置设备发送的 Join-Prune 报文的最大长度，取值范围为 100～8 100 的整数个字节。如果通过此命令配置的报文长度大于接口 MTU 值，则实际报文发送最大长度为接口 MTU 值 缺省情况下，PIM-SM 发送的 Join-Prune 报文的最大长度是 8 100 字节，可用 **undo join-prune max-packet-length** 命令恢复 PIM-SM 发送的 Join-Prune 报文长度为缺省值	
	12	**join-prune periodic-messages queue-size** *queue-size* 例如：[HUAWEI-pim] **join-prune periodic-messages queue-size** 50	配置设备每秒发送 Join-Prune 报文中包含的表项数目，取值范围为 16～4 096 的整数 缺省情况下，PIM-SM 每秒发送 Join-Prune 报文中包含 1 020 个表项，可用 **undo join-prune periodic-messages queue-size** 命令恢复 PIM-SM 每秒发送周期性 Join-Prune 报文中包含的表项数目为缺省值	
	13	**join-prune triggered-message-cache disable** 例如：[HUAWEI-pim]**join-prune triggered-message-cache disable**	去使能实时触发的 Join-Prune 报文打包功能。打包发送 Join-Prune 报文比发送大量 Join-Prune 小报文效率高，因此，设备缺省是将触发性 PIM Join-Prune 小报文打包发送的。若不需要此打包发送机制时，可以通过执行此命令去使能打包功能 缺省情况下，使能实时触发的 Join-Prune 报文打包功能，可用 **undo join-prune triggered-message- cache disable** 命令使能 Join-Prune 报文打包发送功能	
调整剪枝延迟时间	14	**hello-option lan-delay** *interval* 例如：[HUAWEI-pim] **hello-option lan-delay** 1000	全局配置发送剪枝报文的延迟时间，取值范围为 1～32 767 的整数毫秒 缺省情况下，共享网段上传输 Prune 报文的延迟时间是 500ms，可用 **undo hello-option lan-delay** 命令恢复全局剪枝报文延迟时间为缺省值	全局配置
	15	**hello-option override-interval** *interval* 例如：[HUAWEI-pim] **hello-option override-interval** 2000	配置 Hello 报文中携带的否决剪枝的时间间隔，取值范围为 1～65 535 的整数毫秒 缺省情况下，Hello 报文中携带的否决剪枝的时间间隔是 2 500ms，可用 **undo hello-option override-interval** 命令恢复全局否决剪枝的时间间隔为缺省值	
	16	**quit** 例如：[HUAWEI-pim] **quit**	退出发 PIM 视图，返回系统视图	在接口上配置

（续表）

配置任务	步骤	命令	说明	
调整剪枝延迟时间	17	**interface** *interface-type interface-number* 例如：[HUAWEI] **interface vlanif 100**	（可选）键入要配置剪枝延迟时间的 PIM-DM VLAN 接口或者 Loopback 接口，进入接口视图	在接口上配置
	18	**pim hello-option lan-delay** *interval* 例如：[HUAWEI-Vlanif100] **pim hello-option lan-delay** 1000	（可选）在以上接口上配置在 LAN 内传输消息的延迟时间，取值范围为 1～32 767 的整数毫秒 缺省情况下，共享网段上传输 Prune 报文的延迟时间是 500ms，可用 **undo pim hello-option lan-delay** 命令恢复接口上枝报文延迟时间为缺省值	
	19	**pim hello-option override-interval** *interval* 例如：[HUAWEI-Vlanif100] **pim hello-option override-interval** 2000	（可选）在以上接口上配置 Hello 报文中携带的否决剪枝的时间间隔，取值范围为 1～65 535 的整数毫秒 缺省情况下，Hello 报文中携带的否决剪枝的时间间隔是 2 500ms，可用 **undo pim hello-option override-interval** 命令恢复接口上否决剪枝的时间间隔为缺省值	

13.2.6 调整嫁接控制参数

为使被剪枝网段快速恢复转发，设备会向上游发送 Graft 报文请求恢复组播数据转发，并同时在发送接口启动定时器（缺省为 **3s** ）。超时后，如果设备仍没有接收到组播数据，会重新向上游发送 Graft 报文。通过调整嫁接控制参数，可以控制组播数据报文的转发来支持不同转发场景。同样，在调整嫁接控制参数之前，需要完成PIM-DM基本功能配置。

调整嫁接控制参数的方法很简单，就是在对应的VLAN接口或者Loopback接口视图下使用**pim timer graft-retry interval** 命令在接口上配置重传Graft报文的时间间隔，取值范围是 1～65 535的整数秒。缺省情况下，接口上重传Graft报文的时间间隔是 3s，可用**undo pim timer graft-retry** 命令恢复重传Graft报文的时间间隔为缺省值。

【示例 】在VLANIF100接口上配置重传Graft报文的时间间隔为80s。

<HUAWEI> system-view

[HUAWEI] multicast routing-enable

[HUAWEI] interface vlanif 100

[HUAWEI-Vlanif100] pim timer graft-retry 80

13.2.7 调整状态刷新控制参数

为防止被剪枝接口因为剪枝状态超时而恢复转发，PIM-DM网络启用了状态刷新功能，通过在与组播源直连的第一跳**PIM**设备 上周期性地扩散发送State-Refresh报文，刷新接口剪枝定时器，维持SPT树。同样，在调整状态刷新控制参数之前，需要完成PIM-DM基本功能配置。

在整个调整状态刷新控制参数配置过程中，可以配置以下几个方面：禁止状态刷新报文的转发、状态刷新报文的TTL值、状态刷新报文的时间控制参数。配置时无先后顺 序，用户可根据实际需要进行调整。

1．禁止状态刷新报文的转发

缺省情况下，为了避免下游一直没有组播需求的被剪枝接口因为超时而恢复转发，与组播源S直连的PIM设备会触发发送（S，G）状态刷新报文。该报文会逐跳向下游扩散，刷新所有PIM设备上的剪枝定时器。这样没有转发需求的接口将一直处于抑制转发状态。

如果希望组播数据每一次“扩散-剪枝”时都能在全网扩散，不需要通过设备转发状态刷新报文来抑制被剪枝接口转发组播数据，可在接口上禁止此功能。但状态刷新机制能够很好地减少网络资源浪费，一般情况下不建议禁止接口的状态刷新报文的收发能力。

2．调整状态刷新报文的时间控制参数

与组播源直连的第一跳PIM设备会周期性（缺省值为**60s** ）地向下游发送状态刷新报文。由于状态刷新报文扩散发送，设备很有可能在短时间内收到重复的状态刷新报文。为了避免这种情况发生，设备在收到针对某（S，G）的状态刷新报文后，就会启动定时器（缺省值为**30s** ），时间设为该报文的抑制时间。在定时器超时前，如果收到相同的状态刷新报文，就会直接丢弃。

3．配置状态刷新报文的TTL值

设备在收到状态刷新报文后，会将状态刷新报文的TTL值（缺省值为**255** ）减1，然后继续向下游扩散转发来刷新下游设备的剪枝定时器，直至状态刷新报文的TTL值为0。当网络规模很小而TTL值很大时，会造成状态刷新报文在网络中循环传递。因此，为了有效控制刷新报文的传递范围，需要根据网络规模大小配置合适的TTL值。但因为状态刷新报文是由与组播源直连的第一跳**PIM**设备触发发送，所以状态刷新报文的**TTL**值只在该设备上配置有效 。

以上三方面配置任务的具体配置步骤如表13-11所示。

表13-11 调整状态刷新控制参数的配置步骤

配置任务	步骤	命令	说明
公共配置	1	**system-view** 例如：<HUAWEI> **system-view**	进入系统视图
禁止状态刷新报文的转发	2	**interface** *interface-type interface-number* 例如：[HUAWEI] **interface vlanif 100**	键入要禁止状态刷新报文转发的PIM-DM VLAN接口或者Loopback接口，进入接口视图
	3	**undo pim state-refresh-capable** 例如：[HUAWEI-Vlanif100]**undo pim state-refresh-capable**	禁止状态刷新报文的转发。禁止PIM-DM状态刷新后，接口在剪枝定时器超时后开始转发组播数据，不希望接受此数据的下游交换机发送Prune报文进行剪枝。该过程周期性重复，占用较多的网络资源。因此，使能PIM-DM状态刷新，可以在一定程度上优化网络流量 缺省情况下，使能PIM-DM状态刷新，可在接口上执行命令 **pim state-refresh-capable** 重新启用此功能
调整状态刷新报文的时间控制参数	4	**quit** 例如：[HUAWEI-Vlanif100] **quit**	退出接口视图，返回系统视图

（续表）

配置任务	步骤	命令	说明
调整状态刷新报文的时间控制参数	5	**pim** 例如：[HUAWEI] **pim**	进入PIM视图
	6	**state-refresh-interval** *interval* 例如：[HUAWEI-pim] **state-refresh-interval** 100	在与组播源直接相连的第一跳的PIM设备上配置状态刷新报文的发送周期，取值范围为1～255的整数秒 【说明】PIM-DM网络中，设备会周期性地发送状态刷新报文，刷新下游设备启动剪枝定时器的超时时间，使没有组播需求的接口一直处于剪枝状态。执行此命令可设置状态刷新报文的发送周期 缺省情况下，发送PIM状态刷新报文的时间间隔是60s，可用 **undo state-refresh-interval** 命令恢复刷新时间间隔为缺省值
	7	**state-refresh-rate-limit** *interval* 例如：[HUAWEI-pim] **state-refresh-rate-limit** 200	在所有设备上配置相同的状态刷新报文抑制时间，即配置其他PIM设备接收新PIM状态刷新消息前必须经过的最小时间长度，取值范围为1～65 535整数秒 缺省情况下，接收新PIM状态刷新消息前必须经过的最小时间是30s，可用 **undo state-refresh-rate-limit** 命令恢复为缺省值
配置状态刷新报文的TTL值	8	**state-refresh-ttl** *ttl-value* 例如：[HUAWEI-pim] **state-refresh-ttl** 10	在与组播源直接相连的第一跳的PIM设备上配置发送PIM状态刷新消息的TTL值，取值范围是1～255的整数 缺省情况下，发送PIM状态刷新消息的TTL值是255，可用 **undo state-refresh-ttl** 命令恢复TTL值为缺省值

13.2.8 调整断言控制参数

当一个网段内 有多个相连的PIM设备通过RPF检查后向该网段转发组播数据时，则需要通过断言竞选来保证只有一个PIM设备向该网段转发组播数据。在竞选中落败的PIM设备会抑制相应下游接口向该网段转发组播数据。但是这种竞选失败的状态只会保持一段时间，这段时间称为Assert报文的保持时间。超时后，落选的设备会重新恢复转发组播数据从而触发新一轮的竞选。

当设备从下游接口接收到组播数据时，说明该网段中还存在其他的上游设备。设备从该接口发出Assert报文，参与竞选唯一上游。可调整断言Assert报文保持时间（缺省值为**180s**），可在全局PIM视图下或接口视图下配置。如果同时配置，则接口视图上的配置生效。但在调整前需要先完成PIM-DM基本功能配置。

在全局的配置方法是在PIM视图下使用holdtime assert interval命令配置Assert报文的保持时间；在接口上

的配置方法是在具体的VLAN接口或者Loopback接口视图下使用pimholdtime assert interval命令配置Assert报文的保持时间，取值范围均为7～65 535的整数秒。缺省情况下，交换机上所有PIM接口保持Assert状态的超时时间是180s，分别可用undo holdtime assert和undo pim holdtime assert命令恢复超时时间为缺省值。

【示例 1】在PIM视图中全局配置交换机保持Assert状态的超时时间为100s。

```
<HUAWEI> system-view
[HUAWEI] multicast routing-enable
[HUAWEI] pim
[HUAWEI-pim] holdtime assert100
```

【示例 2】配置VLANIF100接口的保持Assert状态的超时时间为100s。

```
<HUAWEI>system-view
[HUAWEI] multicast routing-enable
[HUAWEI] interfacevlanif 100
[HUAWEI-Vlanif100] pim holdtime assert 100
```

13.2.9 配置PIM Silent

使能 PIM Silent功能的接口会禁止接收和转发任何 PIM协议报文，删除该接口上的所有PIM邻居以及PIM状态机。但是，该接口上的IGMP功能不受影响。该功能仅适用于与用户主机网段直连的**PIM**设备接口，且该用户网段只与这一台**PIM**设备相连 。

配置接口为PIM Silent状态的方法很简单，只需在对应的VLAN接口或者Loopback接口视图下配置pim silent命令，使能PIM Silent功能即可。

13.2.10 PIM-DM管理

在 PIM 域内的所有设备上都使能了 PIM-DM 及相关功能之后，可以通过一系列的display任意视图命令查看PIM接口、PIM邻居和PIM路由表，以及其他功能（如剪枝、嫁接、断言等）参数配置信息，以验证配置结果，使用 reset 用户视图命令可以清除指定下游接口的PIM路由表项。

（1）使用display pim interface [interface-type interface-number |up |down] [verbose] 命令查看所有或者指定接口，或者所有状态为Up或者Dwon的接口的PIM信息。

（2）使用 display pimneighbor [neighbor-address | interface interface-type interface- number | verbose] *命令查看指定地址或者（和）指定接口上的详细（选择 verbose可选项时）或者摘要PIM邻居信息。

（3）使用 display pimrouting-table [group-address [mask {group-mask-length | group- mask }] | source-address [mask { source-mask-length | source-mask }] | incoming- interface { interface-type interface-number | register } |outgoing-interface {include | exclude | match } { interface-type interface-number |register |none } |mode {dm | sm | ssm } | flags flag- value|fsm] * [outgoing-interface-number [number]] 命令查看符合条件的PIM路由表详细信息。

（4）使用display pim routing-tablebrief [group-address [mask {group-mask-length |group-mask }] | source-address [mask {source-mask-length | source-mask }] | incoming- interface { interface-type interface-number | register }] *命令查看符合条件的PIM路由表摘要信息。

（5）使用display pim control-message counters [message-type { assert | graft | graft- ack |hello | join-prune | state-refresh |bsr } | interface interface-type interface-number] *命令查看发送和接收PIM控制报文的数目信息。

（6）使用display pim grafts命令查看未确认的PIM-DM嫁接信息。

（7）使用display pim control-message counters [message-type { assert | graft | graft- ack |hello | join-prune | state-refresh |bsr } | interface interface-type interface-number] *命 令查看发送和接收PIM控制报文的数目信息。

（8）使用display pim invalid-packet [interface interface-type interface-number |message- type { assert | graft | graft-afk |hello |join-prune | state-refresh }] *命令查看设备接收到的无效PIM报文的统计信息。

（9）执行reset pimrouting-tablegroupgroup-addressmask {group-mask-length |group- mask } source source-address interface interface-type interface-number 命令清除指定 PIM表项的指定下游接口的PIM状态。

13.2.11 PIM-DM基本功能配置示例

图13-5所示为一个用户比较密集的小型网络，用户主机HostA、HostB希望能够接收到Source发送的组播数据。

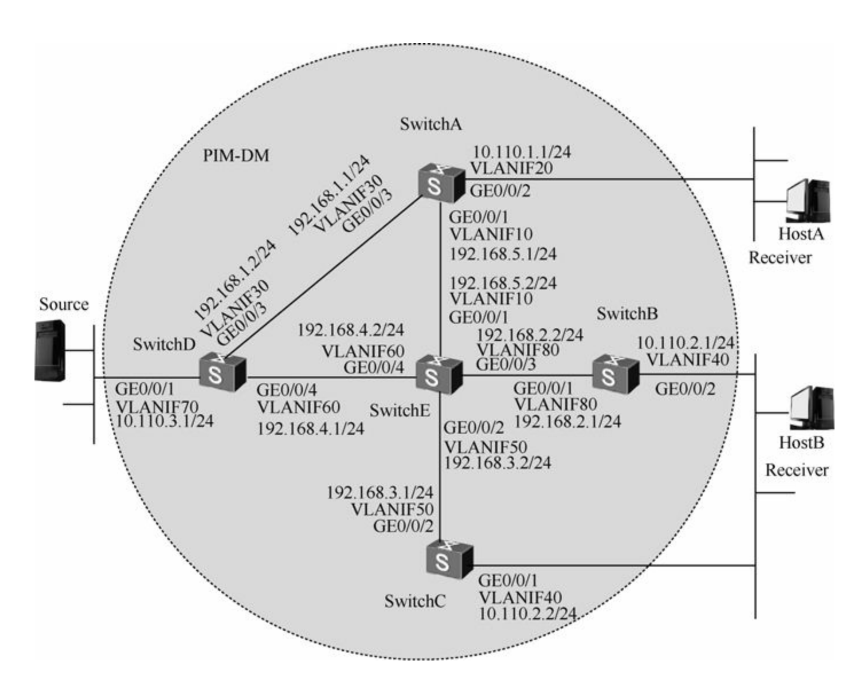


图13-5 PIM-DM基本功能配置示例拓扑结构

1．基本配置思路分析

本示例的要求很简单，就是要求在这样一个密集型小型组播网络中，各组播组成员可以接收到组播源发来的组播数据。所以可以使用PIM-DM协议为网络中的用户主机提供组播服务，使得加入同一组播组的所有用户主机能够接收组播源发往该组的组播数据。具体配置任务如下。

（1）按照图中标注配置交换机各VLAN接口IP地址和单播路由协议，因为组播域内路由协议PIM依赖单播路由协议，单播路由是组播协议正常工作的基础。

（2）在所有提供组播服务的交换机上使能组播路由功能，并在各VLAN接口上使能PIM-DM功能。使能PIM-DM功能之后才能配置PIM-DM的其他功能。

（3）在与主机侧相连的交换机VLAN接口上使能IGMP，用于维护组成员关系。但是，如果用户主机侧需同时配置**PIM-DM**和**IGMP**，必须先使能**PIM-DM**，再使能**IGMP** 。

2．具体配置步骤

（1）配置各交换机接口的IP地址和掩码，各交换机间采用OSPF进行互连，确保网络中各交换机间能够在网络层互通，并且之间能够借助单播路由协议实现动态路由更新。因为SwitchA、SwitchB、SwitchC、SwitchD和SwitchE这五台交换机上的配置一样，下面仅以SwitchA上的配置为例进行介绍。

```
[SwitchA] vlan batch 10 20 30    #---批量创建VLAN 10、VLAN 20和VLAN 30
[SwitchA] interfacevlanif 10
[SwitchA-Vlanif10] ip address 192.168.5.1 24    #---为VLANIF10接口配置 IP地址
[SwitchA-Vlanif10] quit
[SwitchA] interface vlanif 20
[SwitchA-Vlanif20] ip address 10.110.1.1 24
[SwitchA-Vlanif20] quit
[SwitchA] interfacevlanif 30
[SwitchA-Vlanif30] ip address 192.168.1.1 24
[SwitchA-Vlanif30] quit
[SwitchA] interfacegigabitethernet 0/0/1
[SwitchA-GigabitEthernet0/0/1] port hybrid tagged vlan 10    #---把GE0/0/1接口以带标签方式加入VLAN 10
[SwitchA-GigabitEthernet0/0/1] port hybrid pvid vlan 10    #---配置GE0/0/1接口的缺省VLAN为VLAN 10
[SwitchA-GigabitEthernet0/0/1] quit
[SwitchA] interface gigabitethernet 0/0/2
[SwitchA-GigabitEthernet0/0/2] port hybrid tagged vlan 20
[SwitchA-GigabitEthernet0/0/2] port hybrid pvid vlan 20
[SwitchA-GigabitEthernet0/0/2] quit
[SwitchA] interfacegigabitethernet 0/0/3
[SwitchA-GigabitEthernet0/0/3] port hybrid tagged vlan 30
[SwitchA-GigabitEthernet0/0/3] port hybrid pvid vlan 30
[SwitchA-GigabitEthernet0/0/3] quit
[SwitchA] ospf
[SwitchA-ospf-1] area 0    #---进入OSPF骨干区域
[SwitchA-ospf-1-area-0.0.0.0] network192.168.5.0 0.0.0.255    #---宣告192.168.5.0/24网络
[SwitchA-ospf-1-area-0.0.0.0] network192.168.1.0 0.0.0.255
[SwitchA-ospf-1-area-0.0.0.0] network10.110.1.0 0.0.0.255
```

（2）在所有交换机使能组播路由功能，并在各 VLAN 接口上使能 PIM-DM 功能。同样因为SwitchA、SwitchB、SwitchC、SwitchD和SwitchE上的配置方法一样，所以下面也仅以SwitchA上的配置为例进行介绍。

```
[SwitchA] multicast routing-enable
[SwitchA] interface vlanif 10
[SwitchA-Vlanif10] pim dm
[SwitchA-Vlanif10] quit
[SwitchA] interfacevlanif 20
[SwitchA-Vlanif20] pim dm
[SwitchA-Vlanif20] quit
```

[SwitchA] interfacevlanif 30

[SwitchA-Vlanif30] pim dm

[SwitchA-Vlanif30] quit

（3）在SwitchA连接用户主机的接口上使能IGMP功能。SwitchB和SwitchC上的配置过程与SwitchA上的配置相似，配置过程略。

[SwitchA] interface vlanif 20

[SwitchA-Vlanif20] igmp enable

配置好后，可以使用display pim interface命令查看接口上PIM的配置和运行情况，以验证配置结果。例如SwitchC上PIM的显示信息如下，表明接口上的PIM协议已经运行。

<SwitchC>display pim interface

VPN-Instance: public net

Interface	State	NbrCnt	HelloInt	DR-Pri	DR-Address
Vlanif40	up	0	30	1	10.110.2.2 (local)
Vlanif50	up	1	30	1	192.168.3.1 (local)

可使用 display pim routing-table命令查看 PIM协议组播路由表。假设组播源（10.110.3.100/24）已向组播组（225.1.1.1/24）发送信息，而 HostA、HostB 都加入了组播组（225.1.1.1/24）。下面是 SwitchA 交换机上的显示信息，可以看到这个组播组。

```
[SwitchA] display pim routing-table
VPN-Instance: public net
Total 0 (*, G) entry; 1 (S, G) entry
(10.110.3.100, 225.1.1.1)
    Protocol:pim-dm, Flag: ACT
    UpTime: 00:00:29
    Upstream interface: vlanif10
      Upstream neighbor: 192.168.1.2
      RPF prime neighbor: 192.168.1.2
  Downstream interface(s) information:
    Total number of downstreams: 1
      1: vlanif20
        Protocol:pim-dm, UpTime: 00:00:29, Expires:-
```

13.3 PIM-SM（IPv4）配置与管理

PIM-SM属于稀疏模式的域内组播路由协议。它与PIM-DM不同的是，PIM-SM不会将组播数据扩散到全网，而只将组播数据传输到有组成员的网络，一般用于规模较大、组成员分布稀疏的组播网络，且**PIM-SAM**同时适用于**ASM**模型和**SSM**模型（**PIM-DM**仅适用于**ASM**模型）。

与PIM-DM ASM以组播源为转发中心和SPT路径起点不同的是，在PIM-SM ASM模型中，RP是网络的转发中心和SPT路径的起点，网络中所有PIM路由器都知道RP的位置。当网络中出现组成员时，连接组成员的最后一跳PIM路由器向RP方向发送Join信息，然后沿着到达RP单播路由逆向路径向组成员端传递，并逐跳创建（*，G）表项，生成一棵以RP为根的RPT，所以它采用**"拉"**（**Pull**）模式来转发组播报文，即由组

成员主动申请。当网络中出现活跃的组播源时，第一跳 PIM 路由器将组播信息封装在Register报文中发往RP，在RP上创建（S，G）表项，注册源信息。然后，RP会将注册信息中的组播信息解封装，沿着RPT转发到有组成员的网段。有关**PIM-SM SSM**的 工作原理参见本书第**12**章**12.3.3**节 。

如果当前RP、RPT负担较重，可通过以下SPT切换方式减轻压力。

（1）RP向组播源方向发送Join信息，构建“源-RP”的SPT。

（2）组成员端DR向组播源方向发送Join信息，构建“源-组成员”的SPT。

图13-6所示为一个SPT切换示例，最终所有组播成员都采用以RP为核心的SPT路径进行数据转发（如左图所示，SwitchB为RP），后面SwitchC经过向源方向的SPT切换，最终为 HostA 生成了另一条不经过 RP 的新路径（如右图所示），减轻了 RP的压力。

在SSM模型中，网络用户能够预先知道组播源的具体位置。因此用户在加入组播组时，可以明确指定从哪些源接收信息。所以在SSM模型中，与ASM模型的最大区别就是 **SSM**模型中无需维护 **RP**、无需构建 **RPT**（汇集点树）、无需注册组播源，可以直接在源与组成员之间建立**SPT** 。组成员端DR了解到用户的需求后，直接向组播源方向发送Join信息。Join信息按照到达组播源的单播路由路径逐跳向上传输，在源与组成员之间建立 SPT。有关**PIM-SM SSM**的工作原理参见本书第 **12**章**12.3.4**节 。

如图13-7所示，HostA、HostB都已经加入了组播组G，HostA需要接收S1的组播数据，HostB需要接收S2的组播数据，各自的组成员端DR向各自源方向发送Join信息，构建SPT。

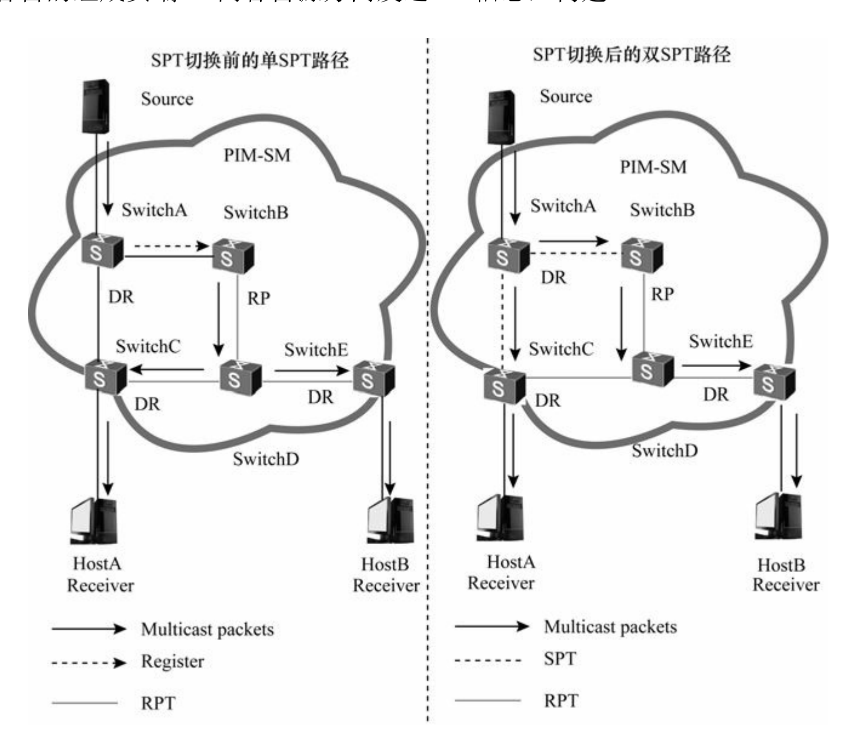


图13-6 ASM模型SPT切换前后对比示意图

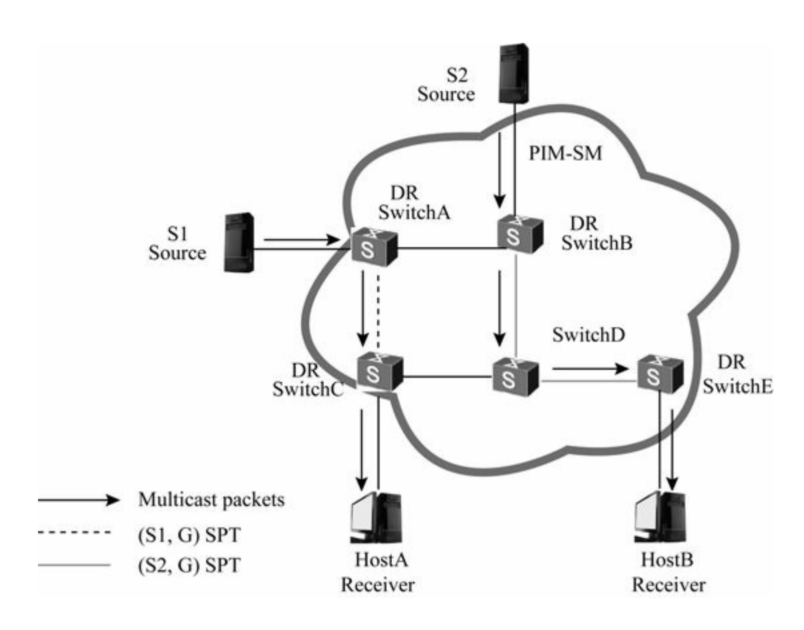


图13-7 SSM模型中SPT构建示意图

13.3.1 PIM-SM（IPv4）特性的产品支持

在 PIM-SM（IPv4）中支持的产品特性有 PIM-SM for ASM、PIM-SM for SSM、调整 PIM-SM控制参数、PIM BFD、PIM GR、PIM Silent。但 **S2700/3700**系列交换机不支持 **PIM GR**、**PIM Silent** 特性。下面分别对这些 PIM-SM特性予以简单介绍。

1．PIM-SM for ASM

ASM模型中，设备支持的PIM-SM基本特性如下。

● 静态/动态RP（**S2700/3700**系列交换机不支持动态**RP** ）。

● BSR管理域（**S2700/3700**系列交换机不支持 ）。

● SPT切换条件的配置。

● 源注册控制参数的调整。

● C-RP/C-BSR参数的调整（**S2700/3700**系列交换机不支持 ）。

2．PIM-SM for SSM

在SSM模型中，设备支持通过配置SSM组策略来指定SSM组地址范围。

3．调整PIM-SM控制参数

在配置ASM或SSM的PIM-SM基本功能后，通过设备提供的缺省值，PIM-SM域就可以正常工作，将组播源发出的组播数据分发到组成员网段。同时也可以根据实际需要，对表13-12所示的PIM-SM控制参数进行调整，多数参数与13.2.1节表13-6中所列的控制参数一样，只是在**PIM-SM**网络中多了**“DR**竞选控制参数**”**和**“**组成员加入控制参数**”**，而少了**PIM-DM**网络中的**“**嫁接控制参数**”**和**“**状态刷新控制参数**”** 。

表13-12 可调整的PIM-SM控制参数

参数	说明
组播源控制参数	设备可以基于组播源来控制组播报文的转发。这样做一方面有助于数据流量控制，另一方面可以限定下游组播组成员能够获得的信息，以提高安全性
邻居控制参数	设备间通过交互 Hello 报文建立 PIM 邻居关系，协商各类控制参数，这样一来可以基于 Hello 报文来控制邻居间的关系。同时可以配置灵活的邻居控制策略，防止非法邻居关系的建立，以提高安全性
DR 竞选控制参数	无论是与组播源相连的网络，还是与组成员相连的网络，都需要选举 DR，由 DR 负责转发组播源或组成员发来的组播报文。可以根据实际需要调整设备上 DR 的优先级；或者由 DR 变为非 DR 时的延迟切换时间
加入和剪枝控制参数	设备向上游发送 Join 信息请求转发组播数据，发送 Prune 信息请求停止转发组播数据。可根据实际需要调整加入或剪枝过程的控制参数，达到控制组播报文转发的目的
断言控制参数	当设备从下游接口接收到组播数据时，说明该网段中还存在其他的上游设备。设备可从该接口发出 Assert 报文，参与竞选唯一上游的转发者。可根据实际需要调整发送断言的间隔，来控制断言竞选的周期

4．PIM BFD

正常情况下，如果共享网段上的当前DR（每个网段都会选举一个**DR**的 ）出现故障，其他 PIM 邻居会等到邻居关系超时才触发新一轮的 DR 竞选过程，这样组播数据传输中断的时间会比较长（不小于邻居关系的超时时间，通常是秒级）。当使能BFD功能后，其故障检测可以达到毫秒级，可大大缩短因DR出现故障而使组播数据传输中断的时间。因为BFD功能可快速地检测共享网段上PIM邻居的状态，当检测到对端故障后立即上报 PIM 模块，然后立即触发新一轮的 DR 竞选过程，而不是等到邻居关系超时。

PIM BFD 功能也适用于共享网段上 Assert（断言）竞选过程，以快速检测 Assert Winner接口故障。

5．PIM GR

在堆叠系统中有时候会进行主备倒换。如果堆叠系统中有组播业务在运行，在主交换机和备交换机进行倒换后，新的主交换机将重新学习PIM路由表及组播转发表，这样会造成组播流量在学习期间的断流。在堆叠系统中配置 PIM GR（Graceful Restart）功能后，主交换机会向备交换机备份PIM路由表项、组播转发表以及需要向上游发送的 Join/Prune 信息。这样在主备倒换后，新的主交换机就可以主动快速地向上游发送 Join 信息，维持上游的加入状态。同时， PIM 协议向所有使能了PIM-SM的交换机发送携带新Generation ID的Hello报文，当下游交换机发现其邻居的Generation ID发生了变化，便向邻居发送 Join/Prune报文以帮助其重新建立路由表项，从而保证转发平面组播数据的不间断转发。但**S2700/3700**系列交换机不支持该特性。

另外，在堆叠系统中配置PIM GR功能后，如果网络中使用的是动态RP，在网络中的DR或次DR收到Generation ID改变的Hello报文后，会向发生主备倒换的堆叠系统单播发送Bootstrap报文（自举报文），堆叠系统可从该自举报文中学习并恢复RP信息。如果堆叠系统未能从自举报文中学习到网络中的 RP 信息，则还可从下游发送的 Join/Prune报文中获取RP信息，重新创建组播路由表。

6．PIM Silent

PIM Silent功能用来禁止用户侧组播设备接口接收和转发任何PIM协议报文，并删除该接口上的所有PIM邻居以及PIM状态机，以防止恶意的PIM Hello报文攻击。详情可参见本章13.2.1节相关介绍。

与PIM-SM相关的功能和参数的缺省配置如表13-13所示。

表13-13 PIM-SM（IPv4）相关功能和参数的缺省配置

参数	缺省值
组播路由功能	未使能
PIM-SM	未使能
静态 RP 地址	未配置
C-RP 接口	未指定
C-BSR 接口	未指定
DR 优先级	1
SPT 切换条件	RP 或组成员端 DR 接收到第一个组播数据报文时就进行 SPT 切换
SSM 组地址范围	232.0.0.0/8
PIM BFD	未使能
PIM GR	未使能
PIM Silent	未使能

13.3.2 ASM模型PIM-SM的配置任务

通过配置ASM模型的PIM-SM，可为用户主机提供任意源组播服务，加入同一组播组的用户主机都能收到任意源发往该组的组播数据。与PIM-DM网络一样，在配置ASM模型的PIM-SM之前也需配置单播路由协议，保证网络内单播路由畅通。

ASM模型的PIM-SM的配置任务如下，其中“使能PIM-SM”和“配置RP”为必选配置任务，其他的均为可选配置任务 。

1．使能PIM-SM

在PIM-SM网络中，在使能了组播路由功能后，首先要使能的就是PIM-SM功能，但设备上不能同时使能**PIM-DM**和 **PIM-SM** 。建议将处于PIM-SM 域内的所有接口都使能PIM-SM，以确保与相连PIM设备都能建立邻居关系。

如果接口上需要同时使能 **PIM-SM** 和 **IGMP**，必须先使能 **PIM-SM**，再使能**IGMP**。

2．配置RP

配置 RP 有手工静态配置和 BSR 机制动态选举两种方式。手工方式静态配置RP 可以避免 C-RP 与 BSR 之间频繁的信息交互而占用带宽。通过 BSR 机制动态选举 RP，可以避免手工配置的繁琐；同时配置多台 C-RP 可以保证组播数据转发的可靠性。

静态**RP**和动态**RP**可同时配置，但此时静态**RP**由于缺省优先级较低而被当作 备份 **RP** 。同时配置时需要确保各组播设备间的 RP 信息一致，否则容易导致网络故障。

表13-14列出了C-BSR、C-RP部分参数的缺省配置。

表13-14 C-BSR、C-RP部分参数的缺省配置

参数	缺省值
C-BSR 优先级	0
C-BSR 携带的哈希掩码长度	30
BSR 报文分片功能	未使能
静态 RP 组播组策略	没有组播组策略，即允许接收任意组地址的组播报文
C-RP 组播组策略	没有组播组策略，即允许接收任意组地址的组播报文
C-RP 优先级	0
C-RP 的宣告报文发送间隔	60s
C-RP 的宣告报文保持时间	150s

3．（可选）配置BSR管理域

为了更有效地管理PIM-SM域，可将PIM-SM域划分为多个BSR（自举路由器）管理域和一个Global域。

每个BSR管理域都维护一个BSR，服务于自己特定地址范围的组播组；Global域也维护一个BSR，为剩余不属于任何BSR管理域的组播组服务。由于一台设备只能加入一个管理域，因此，各个管理域转发组播报文互不干涉；Global可以通过任意管理域内的设备进行报文转发。

BSR管理域可服务的最大组地址范围为239.0.0.0～239.255.255.255。该段地址可重复使用，相当于每个BSR管理域的私有组地址。

4．（可选）配置RPT不向SPT切换

缺省情况下，组成员端DR在接收到第一份组播数据报文后都会向源方向发起SPT切换。如果不希望组成员端DR发起SPT切换，一直用RPT传输组播数据，可配置RPT不向SPT切换功能。

5．（可选）调整注册控制参数

源端DR在收到组播源发送来的组播数据后，会将其封装在注册报文中转发给RP，使得相应组播源可以在RP上注册。注册报文控制参数可在RP和源端DR两个位置进行调整。

在源端DR上可进行如下调整。

（1）配置注册Register抑制时间（缺省为**60s**）。源端DR在收到RP发来的Register-stop（注册停止）报文后，在注册抑制时间内停止向 RP 发送注册报文。超时后，如果源端DR没有收到后续的注册停止报文，则恢复相应注册报文的转发。

（2）配置发送空注册报文时间间隔（缺省为**5s**）。如果注册抑制时间过大或过小，都会影响组播数据的正常转发。通过在抑制期间发空注册报文，可以改善这种影响。

（3）配置仅根据注册报文头来计算校验和（缺省**RP**根据整个注册报文来计算校验和），这样可减少计算校验和的时间，提高注册报文封装组播数据的效率。

（4）配置注册报文的源地址。如果当前源 DR 向 RP 发送的注册报文的源地址对于RP来说不是网络中唯一的IP地址，或者RP上配置了过滤策略将该地址已过滤掉，RP都不会接收到注册报文。此时，通过重新指定合理的源 IP地址，可解决此问题。

在 RP 上可配置过滤注册报文的规则（缺省没有配置过滤策略，即允许接收任意组地址的注册报文），可限定注册报文的地址范围，提高网络安全性。

6．（可选）调整C-RP控制参数

在接口上配置了C-RP（候选RP）后，C-RP会周期性（缺省为**60s**）地向BSR发送Advertisement报文（以下称宣告报文），报文携带该C-RP优先级、该宣告报文的保持时间。BSR在收到该报文后，启动C-RP超时定时器，时间设为宣告报文的保持时间（缺省为**150s**）。在超时前，BSR将宣告报文中携带的C-RP信息汇总成RP-Set信息，封装在自举报文中向PIM域中的所有PIM路由器发送。如果超时后BSR仍没有收到来自某C-RP后续的宣告报文，则认为目前网络中该C-RP失效或不可达。所以**C-RP**发送宣告报文时间间隔必须小于宣告报文的保持时间。

C-RP发送宣告报文时间间隔、C-RP优先级、宣告报文的保持时间都可进行手工配置。有时候为了防止非法C-RP欺骗，还可在BSR上设置合法的C-RP地址范围，只接收该地址范围内C-RP的宣告报文。

7．（可选）调整C-BSR控制参数

BSR由C-BSR（候选BSR）之间自动选举产生。在选举开始时，每个C-BSR都认为自己是本PIM-SM域的BSR，向域内所有PIM路由器发送自举报文。C-BSR在接收到其他C-BSR发来的自举报文后，首先比较二者的优先级；若优先级相同，则再比较二者IP地址，IP地址较大者获胜。获胜者将成为域内的BSR，它会将自己的IP地址和RP-Set信息封装在自举报文中向域内发送。自举报文还携带哈希掩码信息，以备在 C-RP 竞选中进行哈希计算时所需。

BSR周期性（缺省值为**60s**）地发送自举报文，其他的C-BSR收到该报文后会启动超时定时器，时间设

为自举报文的保持时间（缺省值为 **150s** ）；超时后如果没有接收到BSR发来的自举报文，C-BSR之间会触发新一轮的BSR选举过程。所以**BSR**发送自举报文的时间间隔必须要小于自举报文的保持时间 。

C-BSR优先级、BSR哈希掩码、BSR发送自举报文时间间隔、自举报文的保持时间都可进行手工配置。有时候为了防止非法BSR欺骗，还可在接口使能PIM-SM的设备上设置合法的BSR地址范围，只接收该地址范围内BSR的自举报文。

13.3.3 配置ASM模型PIM-SM

上节介绍的 ASM 模型 PIM-SM 网络的各项配置任务的具体配置步骤如表 13-15所示。

表13-15 ASM模型的PIM-SM的配置步骤

配置任务	步骤	命令	说明
公共配置	1	**system-view** 例如：<HUAWEI> **system-view**	进入系统视图
使能PIM-SM	2	**multicast routing-enable** 例如：[HUAWEI] **multicast routing-enable**	全局使能组播路由功能。其他说明参见 13.1.2 节表 13-1 中的第 2 步
	3	**interface** *interface-type interface-number* 例如：[HUAWEI] **interface vlanif** 10	键入要配置 PIM-SM 功能的 VLAN 或者 Loopback 接口，进入接口视图
	4	**pim sm** 例如：[HUAWEI-Vlanif10] **pim sm**	在以上接口上使能 PIM-SM 功能。在接口上使能了 PIM-SM 功能后，交换机才能与相邻的设备建立PIM 邻居，对来自 PIM 邻居的协议报文进行处理 缺省情况下，接口上未使能 PIM-SM，可使用 **undo pim sm** 命令恢复缺省的去使能状态
配置静态 RP	5	**quit** 例如：[HUAWEI-Vlanif10] **quit**	退出接口视图，返回系统视图
	6	**pim** 例如：[HUAWEI] **pim**	进入 PIM 视图
	7	**static-rp** *rp-address* [*basic-acl-number*] [**preferred**] 例如：[HUAWEI-pim] **static-rp** 11.110.0.6 2001 **preferred**	指定静态 RP 地址。当网络内仅有一个 RP 时，可以手工配置静态 RP 而不使用动态 RP，这样可以避免 C-RP 和 BSR 之间频繁的信息交互占用带宽。命令中的参数和选项说明如下。 （1）*rp-address*：指定静态 RP 的 IP 地址 （2）*basic-acl-number*：用于控制所配置的静态 RP 可服务的组播组范围的基本 ACL（过滤的是组播组地址），取值范围为 2 000～2 999 （3）**preferred**：可选项，指定此处配置的静态 RP 优先（**缺省情况下，动态 RP 优先于静态 RP**） 【注意】要在一个 PIM-SM 域内所有的 PIM 设备上都需指定相同的静态 RP 地址，保证静态 RP 正常运行。如果配置的静态 RP 地址是本机某个状态为 UP 的接口地址，本机就作为静态 RP，**但作为静态 RP 的接口不必使能 PIM 协议** 如果没有指定 ACL，则配置的静态 RP 为所有组播组 224.0.0.0/4 服务；如果指定了 ACL，但没有配置规则，则所配置的静态 RP 为所有组 224.0.0.0/4 服务，否则配置的静态 RP 只为能够通过该 ACL 过滤的组播组服务 **重复执行此命令，会配置多个静态 RP，如果存在多个静态 RP 为某个组播组服务的情况，则选择 IP 地址最大的 RP 为该组服务**。当静态 RP 引用的 ACL 规则发生变化时，需要重新为所有组选择静态 RP。对于具有相同 *rp-address* 地址的配置，新配置将覆盖旧配置 缺省情况下，未配置静态 RP，可用 **undo static-rp** *rp-address* 命令删除指定的静态 RP

（续表）

配置任务	步骤	命令	说明	
（可选）配置动态 RP	8	**c-bsr** *interface-type interface-number* [*hash-length* [*priority*]] 例如：[HUAWEI-pim] **c-bsr** vlanif 10	配置 C-BSR，配置动态 RP 首先要配置的是 C-BSR，选举确定 BSR。在一个 PIM-SM 域中，需要配置一个或多个 C-BSR，C-BSR 之间通过自动选举产生 BSR。BSR 负责收集 C-RP 发来的 Advertisement 报文，并将其中 C-RP 的信息汇总成 RP-set 向域内所有设备发送。**建议在组播数据流量汇聚的设备上配置 C-BSR**。命令中的参数说明如下。 （1）*nterface-type interface-number*：指定要配置为 C-BSR 的接口，也只能是 VLAN 接口或者 Loopback 接口 （2）*hash-length*：可选参数，指定该 C-BSR 的哈希掩码长度，取值范围为 0～32 的整数，缺省值为 30。该掩码将被带入哈希函数，用于 RP 竞选 （3）*priority*：可选参数，指定该 C-BSR 的优先级值范围为 0～255 的整数，缺省值为 0。值越大优先级越高 缺省情况下，未配置 C-BSR，可用 **undo c-bsr** 命令恢复缺省配置	配置 C-BSR
	9	**bsm semantic fragmentation** 例如：[HUAWEI-pim]**bsm semantic fragmentation**	（可选）使能 BSR 报文分片功能 【说明】交换机发送 BSR 报文时需要携带网络中所有的 C-RP 信息。当网络中存在大量 C-RP，BSR 报文携带这些 C-RP 信息时，会导致报文长度过大，超过接口 MTU 值，最终可能造成交换机无法正确处理 BSR 报文，从而无法选举出 RP 信息，组播业务也无法正常传输。此时可以使用 BSR 报文分片功能对 BSR 报文进行分片处理，从而保证网络中每台交换机都能学习到一致的 RP 信息，组播分发树能够正确建立。但是必须要保证所有设备都要使能，否则会导致未使能的设备接收到的 RP 信息不完整 缺省情况下，没有使能 BSR 报文分片功能，可用 **undo bsm semantic fragmentation** 命令去使能 BSR 报文分片功能	
	10	**c-rp** *interface-type interface-number* [**group-policy** *basic-acl-number* | **priority** *priority* | **holdtime** *hold-interval* | **advertisement-interval** *adv-interval*] * 例如：[HUAWEI-pim]**c-rp** loopback 0 **group-policy** 2069 **priority 10**	配置交换机向 BSR 通告自己为 C-RP。**建议在组播数据流量汇聚的设备上配置 C-RP**。命令中的参数说明如下。 （1）*interface-type interface-number*：指定要成为 C-RP 的 VLAN 接口或者 Loopback 接口，这样该接口的 IP 地址被通告为 C-RP 地址 （2）**group-policy** *basic-acl-number*：可多选参数，指定用于限定该 C-RP 所服务的组播组的范围的基本 ACL（**过滤的是组播组 IP 地址**），取值范围为 2 000～2 999 （3）**priority** *priority*：可多选参数，指定该 C-RP 的优先级，取值范围为 0～255 的整数，值越大，优先级越低。缺省值为 0	配置 C-RP

（续表）

配置任务	步骤	命令	说明	
（可选）配置动态RP	10	**c-rp** *interface-type interface-number* [**group-policy** *basic-acl-number* | **priority** *priority* | **holdtime** *hold-interval* | **advertisement-interval** *adv-interval*]* 例如：[HUAWEI-pim]**c-rp** loopback 0 **group-policy** 2069 **priority 10**	（4）**holdtime** *hold-interval*：可多选参数，指定BSR等待接收该C-RP发送的Advertisement消息的超时时间，取值范围为1～65 535的整数秒。缺省值为150s （5）**advertisement-interval** *adv-interval*：可多选参数，指定该C-RP发送Advertisement消息的时间间隔，取值范围为1～65 535的整数秒。缺省值为60s 【说明】C-RP竞选RP的规则如下（按顺序比较）。 （1）C-RP接口地址掩码最长者获胜 （2）C-RP优先级较高者获胜 （3）如果优先级相同，则进行Hash计算，结果大者获胜 （4）如果以上都相同，则C-RP的IP地址大者获胜 缺省情况下，交换机未配置C-RP，可用 **undo c-rp** *interface-type interface-number* **undo c-rp** 命令删除指定的C-RP	配置C-RP
	11	**quit** 例如：[HUAWEI-pim] **quit**	（可选）退出PIM视图，返回系统视图	
	12	**Interface** *interface-type interface-number* 例如：[HUAWEI] **interface vlanif** 20	（可选）键入要配置BSR边界的VLAN接口或Loopback接口，进入接口视图。**建议在规划的PIM-SM域的边缘接口配置BSR服务边界**	配置BSR边界
	13	**pim bsr-boundary** 例如：[HUAWEI-Vlanif20] **pim bsr-boundary**	（可选）在以上接口配置BSR服务边界。配置BSR边界后，BSR报文无法通过该边界，主要在划分PIM-SM域时使用，**如果只有一个PIM-SM域，则不用配置** 缺省情况下，未设置PIM-SM域的BSR边界，可用 **undo pim bsr-boundary** 命令取消对应接口上的BSR边界设置	
（可选）配置BSR管理域	14	**quit** 例如：[HUAWEI-Vlanif20] **quit**	退出接口视图，返回系统视图	在PIM-SM域内所有设备上使能BSR管理域功能
	15	**pim** 例如：[HUAWEI] **pim**	进入PIM视图	
	16	**c-bsr admin-scope** 例如：[HUAWEI-pim] **c-bsr admin-scope**	使能交换机的BSR管理域功能 【说明】每个BSR管理域中维护一个BSR，为特定范围239.0.0.0/8网段内的组播组服务，属于该BSR管理域范围内的组播报文无法通过BSR管理域边界。不属于任何BSR管理域的组播组，一律属于Global域的服务范围。Global域中维护一个BSR，为所有剩余的组播组服务，即为组播组地址在239.0.0.0/8范围以外的所有组播组服务 缺省情况下，交换机未使能BSR管理域功能，可用 **undo c-bsr admin-scope** 命令恢复BSR管理域功能缺省的去使能状态	

（续表）

配置任务	步骤	命令	说明	
（可选）配置 BSR 管理域	17	**quit** 例如：[HUAWEI-pim] **quit**	退出 PIM 视图，返回系统视图	在每个 BSR 管理域的边缘接口上配置边界
	18	**Interface** *interface-type interface-number* 例如：[HUAWEI] **interface vlanif** 30	键入 BSR 管理域的边缘接口，进入接口视图	
	19	**multicast boundary** *group-address* { *mask* | *mask-length* } 例如：[HUAWEI-Vlanif30]**multicast boundary** 239.2.0.0 16	在以上 BSR 管理域边缘接口上配置 BSR 管理域的组播地址范围。命令中的参数说明如下。 （1）*group-address*：指定在对应 BSR 管理域中可以转发的组播报文的组播组 IP 地址范围，取值范围是 224.0.1.0～239.255.255.255 （2）*mask*：二选一参数，指定组播组地址的子网掩码，通过它可以确定一个组播组地址范围 （3）*mask-length*：二选一参数，指定组播组地址的子网掩码长度，取值范围是 8～32 的整数。通过它也可以确定一个组播组地址范围 【说明】有时候希望某些组播组的数据在一定范围内转发，比如配置 BSR 管理域时，每个管理域都会有一段特定的组地址为本管理域服务，而组播源发往这些组播组的数据都希望限定在各自的管理域内转发。在接口上配置了针对某些组播组的组播边界之后，指定组播组的组播报文将无法通过该接口进行转发，从而达到了限制转发范围的目的 缺省情况下，任何接口上都没有配置组播转发边界，可用 **undo multicast boundary** { *group-address* { *mask* | *mask-length* } | **all** }命令删除在接口上配置的组播转发边界	

（续表）

配置任务	步骤	命令	说明	
	20	**quit** 例如：[HUAWEI-Vlanif30] **quit**	退出接口视图，返回系统视图	
	21	**pim** 例如：[HUAWEI] **pim**	进入 PIM 视图	
（可选）配置 BSR 管理域	22	**c-bsr group** *group-address* { *mask* | *mask-length* } [**hash-length** *hash-length* | **priority** *priority*] * 例如：[HUAWEI-pim] **c-bsr group** 239.0.0.0 255.0.0.0 **priority** 10	在C-BSR上配置BSR管理域的组播组地址范围。通过在每个管理域的C-BSR上执行该命令，可指定该C-BSR所服务的管理域组地址以及自身的优先级。命令中的*group-address* { *mask* | *mask-length* } 参数参见前面第19步中的对应参数说明，其他两项参数说明如下。 （1）**hash-length** *hash-length*：可多选参数，指定对应组播组在 BSR 管理域中 C-BSR 的哈希掩码长度，取值范围为 0～32 的整数（用于C-BSR选举）。缺省值是30 （2）**priority** *priority*：可多选参数，指定对应组播组在BSR管理域中的C-BSR的优先级（也用于C-BSR选举），取值范围为0～255的整数。值越大，优先级越高。缺省值是0 缺省情况下，未配置C-BSR服务的管理域组地址范围，可用 **undo c-bsr group** *group-address* 命令删除 C-BSR 上配置的组播地址范围	在每个BSR管理域的C-BSR上配置服务的组播组地址范围
	23	**c-bsr global** [**hash-length** *hash-length* | **priority** *priority*] * 例如：[HUAWEI-pim] **c-bsr global priority** 1	配置交换机为Global域中的C-BSR。执行此命令主要用来配置 Global 域中的C-BSR，通过C-BSR竞选产生Global域的BSR。命令中的两个参数与上一步 **c-bsr group** 命令中的对应参数一样，参见即可 缺省情况下，PIM-SM 域中未配置 Global 域的C-BSR，可用 **undo c-bsr global** 命令取消本交换机作为Global域的C-BSR	配置Global域的C-BSR
（可选）配置 RPT 不向 SPT 切换	24	**spt-switch-threshold infinity** 例如：[HUAWEI-pim] **spt-switch-threshold infinity**	在组成员端DR上配置不发起SPT切换 【说明】PIM-SM 组播报文的传输方式为源端 DR 将组播报文封装在注册消息中单播发送至RP，再由RP沿RPT传输到组播组成员。缺省情况下，当RP或者组成员端DR收到第一个组播数据包之后，就会向源发起SPT切换。在组成员端DR配置了此命令后，组成员端DR将永不发起SPT切换 缺省情况下，从 RPT 收到第一个组播数据包后立即进行SPT切换，可用 **undo spt-switch-threshold** 命令禁止切换	

（续表）

配置任务	步骤	命令	说明	
（可选）调整注册控制参数	25	**register-suppression-timeout** *interval* 例如：[HUAWEI-pim] **register-suppression-timeout** 70	在源端 DR 上配置保持注册抑制状态的超时时间，取值范围为 11～3 600 的整数秒 【说明】当交换机接收到从 RP 发来的针对（S，G）项的 Register-Stop 报文，会立刻停止发送封装组播数据的 Register 报文，此时交换机进入注册抑制状态。执行此命令可设置注册抑制状态的超时时间。超时后，源端 DR 将恢复向 RP 发送 Register 报文 缺省情况下，注册抑制状态的超时时间是 60s，可用 **undo register- suppression-timeout** 命令恢复超时时间为缺省值	在源端 DR 上配置
	26	**probe-interval** *interval* 例如：[HUAWEI-pim] **probe-interval** 30	在源端 DR 上配置交换机向 RP 发送 Probe 报文（空注册报文）的时间间隔，取值范围是 1～1 799 的整数秒。但必须小于上一步 **register- suppression-timeout** 值的 1/2 【说明】当组播源侧 DR 收到 RP 发送的 Register-Stop 报文后，组播源端 DR 将会停止发送注册报文并进入注册抑制状态。在注册抑制期间，组播源端 DR 向 RP 周期性发送 Probe 报文以通告组播源仍处于激活状态。注册抑制超时后，组播源端 DR 重新开始发送注册报文 缺省情况下，交换机向 RP 发送 Probe 报文的时间间隔是 5s，可用 **undo probe-interval** 命令恢复时间间隔为缺省值	
	27	**register-header-checksum** 例如：[HUAWEI-pim] **register-header-checksum**	在源端 DR 上配置仅根据 Register 注册报文头信息来计算校验和，未通过校验的 Register 注册报文将被丢弃 【说明】缺省情况下，源端 DR 根据 Register 注册报文全部内容来计算校验和。执行此命令后，源端 DR 仅根据注册报文头来计算校验和，可减少计算校验和的时间，提高注册报文封装组播数据的效率，可用 **undo register-header-checksum** 命令恢复缺省配置	
	28	**register-source** *interface-type interface-number* 例如：[HUAWEI-pim] **register-source** loopback 0	指定源端 DR 发送注册报文的源地址 【说明】如果发送注册报文的源 IP 地址对于 RP 路由器不再是网络中唯一的 IP 地址或者是一个被过滤掉的 IP 地址，那么注册过程就会出现错误，导致网络中出现多余的流量，占用带宽。这时可以通过本命令指定一个源端 DR 上合理接口作为发送注册报文的源 IP 地址，**建议使用源 DR 上 Loopback 接口的 IP 地址** 缺省情况下，不指定源 DR 发送注册报文的源地址，可用 **undo register-source** 命令取消指定的源 DR 发送注册报文的源地址	

（续表）

配置任务	步骤	命令	说明	
（可选）调整注册控制参数	29	**register-policy** *advanced-acl-number* 例如：[HUAWEI-pim] **register-policy** 3001	配置 RP 过滤 Register 注册报文的规则。参数 *advanced-acl-number* 用来指定过滤组播源组地址的高级 ACL 编号，取值范围为 3 000～3 999。在定义 ACL 规则时，通过 **permit** 选项配置设备仅接收指定地址范围的注册报文。如果 ACL 未定义规则，则设备缺省过滤掉所有的注册报文 【说明】为了防止非法注册报文攻击，可以根据报文过滤规则来接受或拒绝和规则匹配的注册报文 缺省情况下，未配置注册报文过滤规则，可用 **undo register-policy** 命令取消注册报文过滤配置	在 RP 上配置
（可选）调整 C-RP 控制参数	30	**c-rp priority** *priority* 例如：[HUAWEI-pim] **c-rp priority** 20	配置 C-RP 的全局性优先级，取值范围是 0～255，优先级数值越大，优先级越低。但重复配置此命令将覆盖原有配置信息。C-RP 竞选 RP 的规则参见以上第 10 步说明 缺省情况下，C-RP 的全局性优先级是 0，可用 **undo c-rp priority** 命令恢复该优先级为缺省值	在 C-RP 上配置宣告报文携带的参数
	31	**c-rp advertisement-interval** *interval* 例如：[HUAWEI-pim] **c-rp advertisement-interval** 30	配置 C-RP 周期性发送 Advertisement 报文的时间间隔，取值范围为 1～65 535 的整数秒 【说明】PIM-SM 域内的所有 C-RP 会周期性地向 BSR 发送携带自身参数的 Advertisement 报文，然后 BSR 将收集到这些 C-RP 信息汇总成 RP-set 向域内所有设备发送。可通过此命令配置 C-RP 向 BSR 发送 Advertisement 报文的时间间隔 缺省情况下，C-RP 发送 Advertisement 报文的时间间隔是 60s，可用 **undo c-rp advertisement-interval** 命令恢复发送时间间隔为缺省值	
	32	**c-rp holdtime** *interval* 例如：[HUAWEI-pim] **c-rp holdtime** 60	配置 C-BSR 等待接收 BSR 发送的 Bootstrap 报文的超时时间，取值范围为 1～214 748 364 的整数秒 【说明】当某 C-BSR 竞选获胜成为 BSR 后，周期性地向网络发送 Bootstrap 报文，报文中携带自己的 IP 地址、RP-Set 信息。Bootstrap 报文的发送间隔为 BS_intervel，可以使用上一步的 **c-bsr interval** 命令配置	

（续表）

配置任务	步骤	命令	说明	
可选）调整C-RP控制参数	32	**c-rp holdtime** *interval* 例如：[HUAWEI-pim] **c-rp holdtime** 60	其他选举落败的C-BSR抑制Bootstrap报文的发送，并启动定时器监视当选BSR。定时器超时时间为Holdtime，可以使用本命令配置。如果收到当选BSR发来的Bootstrap报文，则刷新定时器。落败C-BSR也根据Holdtime刷新BSR的超时时间；如果定时器超时，则认为当选BSR发生故障。落败C-BSR自发执行竞选产生新的BSR，从而确保业务免受中断 缺省情况下，C-BSR等待接收BSR发送的Bootstrap报文的超时时间是130s，可用 **undo c-bsr holdtime** 命令恢复超时时间为缺省值	在C-RP上配置宣告报文携带的参数
	33	**crp-policy** *advanced-acl-number* 例如：[HUAWEI-pim] **crp-policy** 3100	在BSR上配置用来限定合法的C-RP地址范围及其服务的组播组地址范围，使其丢弃来自该地址范围之外的C-RP报文，从而防止C-RP欺骗。参数 *advanced- acl-number* 指定用于定义了针对C-RP地址范围（**作为规则中的源地址**）和其服务组播组地址（**作为规则中的目的地址**）范围的过滤策略的高级ACL，取值范围为3 000～3 999。在定义ACL的规则时，通过 **permit** 选项配置设备仅接收指定地址范围的宣告报文。如果ACL未定义规则，则设备缺省过滤掉所有的宣告报文 【说明】为了防止C-RP欺骗，需要在BSR上配置本命令限定合法的C-RP地址范围以及其服务的组播组地址范围。**由于每个C-BSR都可能成为BSR，因此需要在每个C-BSR上都配置相同的过滤策略** 缺省情况下，C-RP地址范围及其服务的组播组地址范围不受任何限制，即BSR认为接收到的所有C-RP报文都是合法的，可用 **undo crp-policy** 命令恢复缺省配置	在BSR上限定合法的C-RP地址范围
（可选）调整C-BSR控制参数	34	**c-bsr priority** *priority* 例如：[HUAWEI-pim] **c-bsr priority** 100	配置C-BSR的全局优先级（**可能需要在每个C-BSR上配置**），取值范围为0～255的整数。**值越大，优先级越高** 【说明】多个C-BSR与竞选BSR的规则如下。 （1）具有最高优先级的交换机将成为BSR （2）当优先级相同时，IP地址较大者将成为BSR	在C-BSR上配置自举报文携带的参数

（续表）

配置任务	步骤	命令	说明	
（可选）调整C-BSR控制参数	34	**c-bsr priority** *priority* 例如：[HUAWEI-pim] **c-bsr priority** 100	当希望某个C-BSR成为BSR时，可以配置该命令调大该C-BSR的优先级数值 缺省情况下，C-BSR的全局优先级是0，可用 **undo c-bsr priority** 命令恢复该配置参数的缺省值	在C-BSR上配置自举报文携带的参数
	35	**c-bsr hash-length** *hash-length* 例如：[HUAWEI-pim] **c-bsr hash-length** 20	配置C-BSR的全局性哈希掩码长度（**可能需要在每个C-BSR上配置**），取值范围为0～32的整数 【说明】在进行动态RP竞选时，如果C-RP针对特定组的接口地址掩码和优先级都相同，则需要执行哈希函数来选取该组的RP。交换机根据组地址G、C-RP的地址和哈希掩码长度，运用哈希函数，对希望为组G服务且优先级相同的C-RP逐一进行计算，并比较计算结果，计算结果最大者为组播组G提供服务的RP。配置哈希掩码长度主要用来调整哈希计算结果 缺省情况下，C-BSR的全局性哈希掩码长度是30，可用 **undo c-bsr hash-length** 命令恢复该配置参数的缺省值	
	36	**c-bsr holdtime** *interval* 例如：[HUAWEI-pim] **c-bsr holdtime** 100	配置C-BSR等待接收BSR发送的Bootstrap报文的超时时间（**可能需要在每个C-BSR上配置**），取值范围为1～214 748 364的整数秒 【说明】在实际应用中，**属于同一个PIM域的所有C-BSR必须使用相同的BS_interval**（将在下一步介绍）**和Holdtime**。如果配置值不同，有可能导致当选BSR不稳定，从而引发组播故障。有以下注意事项。 （1）如果同时配置了BS_interval和Holdtime，则请务必保证BS_interval小于Holdtime。 （2）如果只配置了其中之一，则使用公式：Holdtime＝2 × BS_interval + 10，计算另一个。如果配置了Holdtime，计算结果小于BS_interval取值范围的最小值时，BS_interval取最小值；如果配置了BS_interval，计算结果大于Holdtime取值范围的最大值时，Holdtime取最大值	

（续表）

配置任务	步骤	命令	说明	
（可选）调整 C-BSR 控制参数	36	**c-bsr holdtime** *interval* 例如：[HUAWEI-pim] **c-bsr holdtime** 100	（3）如果都未配置，则使用缺省值：BS_interval 为 60s，Holdtime 为 130s 缺省情况下，C-BSR 等待接收 BSR 发送的 Bootstrap 报文的超时时间是 130s，可用 **undo c-bsr holdtime** 命令恢复超时时间为缺省值	在 C-BSR 上配置自举报文携带的参数
	37	**c-bsr interval** *interval* 例如：[HUAWEI-pim] **c-bsr interval** 100	配置 C-BSR 发送 Bootstrap 自举报文的间隔时间（**可能需要在每个 C-BSR 上配置**），取值范围为 1～107 374 177 的整数秒 【说明】当某 C-BSR 竞选获胜成为 BSR 后，将周期性地向 PIM-SM 域内发送 Bootstrap 报文，报文中携带自己的 IP 地址、RP-Set 信息。Bootstrap 报文的发送间隔为 BS_intervel，可以使用本命令配置。其他选举落败的 C-BSR 抑制 Bootstrap 报文的发送，并启动定时器监视当选 BSR。定时器超时时间为 Holdtime，可以使用上一步的 **c-bsr holdtime** 命令配置。如果收到当选BSR发来的Bootstrap报文，则刷新定时器；如果定时器超时，则认为当选 BSR 发生故障。落败 C-BSR 自发执行竞选产生新的 BSR，从而确保业务免受中断 缺省情况下，BSR 连续发送 Bootstrap 报文的时间间隔是 60s，可用 **undo c-bsr interval** 命令恢复时间间隔为缺省值	
	38	**bsr-policy** *basic-acl-number* 例如：[HUAWEI-pim] **bsr-policy** 2100	在每个 PIM 设备上限定合法 BSR 地址范围，使交换机丢弃来自该地址范围之外的自举报文，从而防止 BSR 欺骗。参数 *basic-acl-number* 指定用于表定义了针对 BSR 报文源地址范围的过滤策略的基本 ACL，取值范围为 2 000～2 999 在定义 ACL 规则时，通过 **permit** 选项配置设备仅接收指定地址范围的自举报文。如果 ACL 未定义规则，则设备缺省过滤掉所有地址范围的自举报文 缺省情况下，BSR地址范围不受任何限制，即交换机接收到的所有自举报文都认为是有效的，不会丢弃，可用 **undo bsr-policy** 命令恢复缺省配置	在接口使能 PIM-SM 的设备上限定合法的 BSR 地址范围

【示例 1】全局配置地址为11.110.0.6的交换机为静态RP，为ACL2001定义的组播组提供服务，并且启用静态RP优先。

```
<HUAWEI> system-view
[HUAWEI] acl number 2001
[HUAWEI-acl-basic-2001] rule permit source225.1.0.0 0.0.255.255
[HUAWEI-acl-basic-2001] quit
[HUAWEI] multicast routing-enable
[HUAWEI] pim
[HUAWEI-pim] static-rp 11.110.0.6 2001preferred
```

【示例 2】在交换机的VLANIF100上配置C-BSR。

```
<HUAWEI> system-view
[HUAWEI] multicast routing-enable
[HUAWEI] interface vlanif 100
[HUAWEI-Vlanif100] pim sm
[HUAWEI] pim
[HUAWEI-pim] c-bsr vlanif 100
```

【示例 3】全局配置Loopback0接口作为PIM-SM组播域中组播组地址为225.1.0.0/16和226.2.0.0/16的C-RP，并且设置该C-RP的优先级为10。

```
<HUAWEI> system-view
[HUAWEI] acl number2069
[HUAWEI-acl-basic-2069] rule permit source 225.1.0.0 0.0.255.255
[HUAWEI-acl-basic-2069] rule permit source 226.2.0.0 0.0.255.255
[HUAWEI-acl-basic-2069] quit
[HUAWEI] multicast routing-enable
[HUAWEI] pim
[HUAWEI-pim] c-rp loopback 0 group-policy 2069 priority 10
```

【示例 4】 在C-BSR交换机上配置C-RP策略，仅允许1.1.1.1/32的交换机作为C-RP，并且只允许该C-RP为225.1.0.0/16范围的组播组服务（通过高级ACL定义规则）。

```
<HUAWEI>system-view
[HUAWEI] acl number 3100
[HUAWEI-acl-adv-3100] rule permit ip source 1.1.1.1 0destination 225.1.0.0 0.0.255.255
[HUAWEI-acl-adv-3100] quit
[HUAWEI] multicast routing-enable
[HUAWEI] pim
[HUAWEI-pim] crp-policy 3100
```

【示例 5】全局配置合法BSR地址范围是10.1.1.0/24网段。

```
<HUAWEI>system-view
[HUAWEI] acl number2001
[HUAWEI-acl-basic-2001] rule permit source 10.1.1.0 0.0.0.255
[HUAWEI-acl-basic-2001] quit
[HUAWEI] pim
[HUAWEI-pim] bsr-policy 2001
```

13.3.4 配置SSM模型的PIM-SM

SSM模型PIM-SM的配置很简单，主要包括两项配置任务：一是必选的PIM-SM功能，二是可选的SSM组策略配置。通过SSM组策略可用来控制SSM组地址范围。具体如表13-16所示。

表13-16 SSM模型的PIM-SM的配置步骤

步骤	命令	说明
1	**system-view** 例如：<HUAWEI> **system-view**	进入系统视图
2	**multicast routing-enable** 例如：[HUAWEI] **multicast routing-enable**	全局使能组播路由功能。其他说明参见 13.1.2 节表 13-1 中的第 2 步
3	**interface** *interface-type interface-number* 例如：[HUAWEI] **interface vlanif** 10	键入要配置 PIM-DM 功能的 VLAN 或者 Loopback 接口，进入接口视图

（续表）

步骤	命令	说明
4	**pim sm** 例如：[HUAWEI-Vlanif10] **pim sm**	在以上接口上使能 PIM-SM 功能。其他说明参见 13.3.3 节表 13-15 中的第 4 步
5	**quit** 例如：[HUAWEI-Vlanif10] **quit**	退出接口视图，返回系统视图
6	**pim** 例如：[HUAWEI] **pim**	进入 PIM 视图
7	**ssm-policy** *basic-acl-number* 例如：[HUAWEI-pim] **ssm-policy** 2010	（可选）配置 SSM 组播组地址范围，**仅在需要扩展 SSM 组播组地址范围时配置**。参数用来定义 SSM 组播地址范围的基本 ACL，取值范围为 2 000～2 999。**但要确保网络内所有 PIM 设备上配置的 SSM 组地址范围都一致** **缺省情况下，SSM 组范围是 232.0.0.0/8，执行此命令后可以超出这个范围**，所有使能 PIM-SM 协议的接口将会认为属于该范围内的组播组采用了 PIM SSM 模式，可用 **undo ssm-policy** 命令恢复缺省配置

【示例 】配置PIM SSM组播地址范围为 232.1.0.0/16。

```
<HUAWEI> system-view
[HUAWEI] acl number 2000
[HUAWEI-acl-basic-2000] rule permit source232.1.0.0 0.0.255.255
[HUAWEI-acl-basic-2000] quit
[HUAWEI] multicast routing-enable
[HUAWEI] pim
[HUAWEI-pim] ssm-policy 2000
```

13.3.5 PIM-SM其他可选功能及参数配置

本节将要介绍的是 PIM-SM 网络中其他可选功能及参数配置与前面 1.2.2 节中介绍的各项功能及控制参数配置方法基本一样。这些配置选项包括以下几种。

（1）调整组播源控制参数

与在本章13.2.3节介绍的PIM-DM“调整组播源控制参数”配置方法完全一样，参见即可。

（2）调整邻居控制参数

与在本章13.2.4节介绍的PIM-DM“调整邻居控制参数”配置方法基本一样，只是在PIM-SM网络中多了一项“跟踪下游邻居功能”配置，具体配置步骤将在本节后面介绍，其他参数配置参见13.2.4节即可。

（3）调整DR竞选控制参数

这是PIM-DM网络中所没有的，仅在PIM-SM网络需要配置。

设备之间通过交互Hello报文选举DR，主要负责源端或者组成员端的协议报文发送的工作。这里又包括配置DR优先级和配置DR切换延迟两方面。

在“配置DR优先级”方面，组播源或组播成员所在的共享网段，通常同时连接着多台PIM设备。为了争取该网段唯一的组播报文转发权，PIM设备之间就需要通过交互Hello报文进行DR竞选。竞选时，首先比较Hello报文中携带的DR优先级（缺省值为 **1** ），优先级较高者获胜（优先级数值越大，表示优先级越高）；如果DR优先级相同或该网段存在至少一台PIM设备不支持在Hello报文中携带DR优先级，则IP地址较大者获胜。DR优先级在全局PIM视图下和接口视图下都可配置，如果同时配置，接口视图上的配置生效。

在“配置DR切换延迟”方面，有时候由于某些原因，当前共享网段的DR变成非DR，原有向该网段的转发数据的组播表项会被立即删除，这可能会导致短时间内组播数据的断流。此时，可以配置DR切换延迟，并指定延迟时间，原有表项仍然有效直到延迟时间超时。

这部分具体配置步骤也将在本节后面介绍。

（4）调整加入和剪枝控制参数

与在本章13.2.5节介绍的PIM-DM“调整剪枝控制参数”配置方法基本一样，只是在PIM-SM网络中多了一项“Join信息过滤策略”配置，具体将在本节后面介绍，其他参数配置参见13.2.5节即可。

（5）调整断言控制参数

与在本章13.2.8节介绍的PIM-DM“调整断言控制参数”配置方法完全一样，参见即可。

（6）PIM BFD

这部分是 PIM-SM 所特有的功能，用于在检测到对端故障以后立即触发新一轮的DR 竞选过程，而不是等到邻居关系超时，这将在很大程度上缩小组播数据传输的中断时间，提高组播网络的可靠性。具体配置步骤也将在本节后面介绍。

（7）PIM GR

这部分也是PIM-SM网络所特有的功能，用于在设备进行主备倒换时实现快速倒换，保持用户组播流量的正常转发。具体配置步骤也将在本节后面介绍。

（8）PIM Silent

与在本章前面 13.2.9节介绍的 PIM-DM“配置 PIM Silent”配置方法完全一样，参见即可。

下面介绍以上所提到的在IM-SM网络中特有的一些参数和功能的具体配置步骤。

1．配置跟踪下游邻居功能

设备发送Hello报文时，会生成一个Generation ID携带在该报文中。一般Generation ID不会改变，只有设备状态改变，此时Generation ID重新生成才会改变。这时邻居设备在收到Hello报文后，发现Generation ID改变，会立即向该设备发送加入报文以刷新邻居关系。

正常情况下，如果共享网段内有多台设备都准备向同一上游设备发送加入请求，会采用侦听机制来抑制这种相同加入报文的数目，即一台设备在侦听到其他设备的加入报文后，将不会再向该上游 PIM邻居发送加入报文。这时会因Generation ID改变的上游邻居无法刷新与每台下游的邻居关系。配置了“跟踪下游邻居”功能后，设备在侦听到其他设备发送的加入报文时，将不会抑制向相同的上游PIM邻居发送加入报文。

该功能在全局 PIM 视图下和接口视图下都可配置。如果同时配置，接口视图上的配置生效，但必须保证共享网段中的所有设备都使能该功能 。具体配置步骤如表13-17所示。

表13-17 跟踪下游邻居功能的配置步骤

步骤	命令	说明	
1	**system-view** 例如：<HUAWEI> **system-view**	进入系统视图	
2	**pim** 例如：[HUAWEI] **pim**	进入 PIM 视图。可用 **undo pim** 命令清除 PIM 视图下进行的配置，将删除所有 IPv4 PIM 全局配置信息，请慎用	全局使能跟踪下游邻居功能
3	**hello-option neighbor-tracking** 例如：[HUAWEI-pim] **hello-option neighbor-tracking**	全局使能跟踪下游邻居功能 缺省情况下，未使能邻居跟踪功能，可用 **undo hello-option neighbor-tracking** 命令用来恢复缺省配置	
4	**quit** 例如：[HUAWEI-pim] **quit**	退出发 PIM 视图，返回系统视图	在接口上使能跟踪下游邻居功能
5	**interface** *interface-type interface-number* 例如：[HUAWEI] **interface vlanif 100**	键入要配置邻居控制参数的 PIM-DM VLAN 接口或者 Loopback 接口，进入接口视图	
6	**pim hello-option neighbor-tracking** 例如：[HUAWEI-Vlanif100] **pim hello-option neighbor-tracking**	（可选）在以上接口上使能跟踪下游邻居功能 缺省情况下，未使能邻居跟踪功能，可用 **undo pim hello-option neighbor-tracking** 命令用来恢复缺省配置	

2．调整DR竞选控制参数

设备之间通过交互Hello报文选举DR，主要负责源端或者组成员端的协议报文发送的工作。可以配置DR竞选优先级和DR切换延迟功能，无先后顺序，用户可根据实际需要进行调整。

（1）DR优先级。在组播源或组成员所在的共享网段，通常同时连接着多台PIM设备。为了争取该网段唯一的组播报文转发权，PIM设备之间就需要通过交互Hello报文进行DR竞选。竞选时，首先比较Hello报文中携带的DR优先级（缺省值为**1** ），优先级较高者获胜（优先级数值越大，表示优先级越高）；如果DR优先级相同或该网段存在至少一台PIM设备不支持在Hello报文中携带DR优先级，则IP地址较大者获胜。DR优先级在全局PIM视图下和接口视图下都可配置，如果同时配置，接口视图上的配置生效。

（2）配置DR切换延迟。有时候由于某些原因，当前共享网段的DR变成非DR，原有向该网段的转发数据的组播表项会被立即删除，这可能会导致短时间内组播数据的断流。此时可以配置DR切换延迟，并指定延迟时间，原有表项仍然有效直到延迟时间超时。缺省情况下，未配置DR切换延迟功能。

以上两项功能的具体配置步骤如表13-18所示。

表13-18 调整DR竞选控制参数的配置步骤

步骤	命令	说明	
1	**system-view** 例如：<HUAWEI> **system-view**	进入系统视图	
2	**pim** 例如：[HUAWEI] **pim**	进入 PIM 视图。可用 **undo pim** 命令清除 PIM 视图下进行的配置，将删除所有 IPv4 PIM 全局配置信息，请慎用	全局配置 DR 优先级

（续表）

步骤	命令	说明	
3	**hello-option dr-priority** *priority* 例如：[HUAWEI-pim] **hello-option dr-priority** 100	全局配置交换机竞选 DR 的优先级，取值范围为 0～4 294 967 295 的整数 缺省情况下，交换机竞选成为 DR 的优先级是 1，可用 **undo hello-option dr-priority** 命令恢复交换机全局 DR 优先级参数为缺省值	全局配置 DR 优先级
4	**quit** 例如：[HUAWEI-pim] **quit**	退出发 PIM 视图，返回系统视图	在接口上配置 DR 优先级
5	**interface** *interface-type interface-number* 例如：[HUAWEI] **interface vlanif 100**	（可选）键入要配置 DR 优先级的 PIM VLAN 接口或者 Loopback 接口，进入接口视图	
6	**pim hello-option dr-priority** *priority* 例如：[HUAWEI-Vlanif100] **pim hello-option dr-priority** 200	（可选）在以上接口上配置竞选 DR 的优先级，取值范围为 0～4 294 967 295 的整数 缺省情况下，交换机竞选成为 DR 的优先级是 1，**undo hello-option dr-priority** 命令用来恢复对应接口上 DR 优先级为缺省值	
7	**pim timer dr-switch-delay** *interval* 例如：[HUAWEI-Vlanif100] **pim timer dr-switch-delay** 360	在接口上配置 DR 切换延迟，并指定延迟时间，取值范围为 10～3 600 的整数秒。当出接口由 DR 变成非 DR 时，在延迟时间超时之前，出接口继续转发数据 缺省情况下，当出接口由 DR 变为非 DR 时，出接口立即停止转发数据，可用 **undo pim timer dr-switch-delay** 命令取消接口上的 PIM DR 切换延迟功能	配置 DR 切换延迟

3．配置Join信息的过滤策略

有时候为了防止非法用户的加入，还可配置Join信息过滤策略，指定Join-Prune报文中Join信息的合法源地址范围。具体配置方法是在对应的PIM接口视图下使用pimjoin-policy { asm basic-acl-number | ssm advanced-acl-number | advanced-acl-nu mber }命令配置 Join信息过滤策略，限定 Join信息的合法源地址范围。命令中的参数说明如下。

（1）asm basic-acl-number：多选一参数，指定用于过滤在ASM组播组地址的Join信息，取值范围为 2

000～2 999。

（2）ssm advanced-acl-number：多选一参数，指定源地址向组地址在SSM范围内的组播组发送的 Join信息，取值范围为 3 000～3 999。

（3）advanced-acl-number：多选一参数，指定源地址向ASM或者SSM组地址的组播组发送的 Join信息，取值范围为 3 000～3 999。

在定义ACL规则时，通过permit选项配置设备仅接收指定地址范围的Join信息。如果ACL未定义规则，则接口缺省过滤掉Join-Prune报文中所有地址范围的Join信息。

缺省情况下，不过滤 Join-Prune报文中的 Join信息，可用undo pim join-policy命令恢复缺省配置。

【示例 】配置VLANIF100接收组地址范围是225.1.0.0/16的Join信息。

```
<HUAWEI> system-view
[HUAWEI] acl number2001
[HUAWEI-acl-basic-2001] rule permit source 225.1.0.0 0.0.255.255
[HUAWEI-acl-basic-2001] quit
[HUAWEI] multicast routing-enable
[HUAWEI] interface vlanif 100
[HUAWEI-Vlanif100] pim join-policy asm 2001
```

4．配置PIM BFD

启用BFD功能后，可以实现毫秒级的快速故障检测。利用BFD来检测共享网段上PIM邻居的状态，当BFD检测到对端故障以后上报PIM模块，PIM模块立即触发新一轮的DR竞选过程，而不是等到邻居关系超时，这将在很大程度上缩小组播数据传输的中断时间，提高组播网络的可靠性。具体配置步骤如表13-19所示。

表13-19 PIM BFD的配置步骤

<table>
<tr><th>步骤</th><th>命令</th><th>说明</th></tr>
<tr><td>1</td><td>system-view
例如：<HUAWEI> system-view</td><td>进入系统视图</td></tr>
<tr><td>2</td><td>interface interface-type interface-number
例如：[HUAWEI] interface vlanif 10</td><td>键入要配置 BFD 的 PIM 接口，进入接口视图</td></tr>
<tr><td>3</td><td>pim bfd enable
例如：[HUAWEI-Vlanif10] pim bfd enable</td><td>在以上接口上使能 PIM BFD 功能。执行此命令后将使能接口 PIM BFD 功能，可以快速地检测邻居链路故障
缺省情况下，接口没有使能 PIM BFD 功能，undo pim bfd enable 命令取消接口上的 PIM BFD 功能</td></tr>
<tr><td>4</td><td>pim bfd { min-tx-interval tx-value | min-rx-interval rx-value | detect-multiplier multiplier-value } *
例如：[HUAWEI-Vlanif10] pim bfd min-tx-interval 300 min-rx-interval 200 detect-multiplier 10</td><td>调整接口的 PIM BFD 参数。设备上使能了 PIM BFD 功能后，有时候需要调整 PIM BFD 会话的参数，来适应当前链路情况。可通过执行此命令设置 PIM BFD 检测报文的最小发送间隔、最小接收间隔，以及本地检测倍数（检测次数）。但当其他协议配置了同样的 BFD 参数时，PIM BFD 配置的参数可能会受到影响。命令中的参数说明如下。
（1）min-tx-interval tx-value：可多选参数，指定 PIM BFD 报文的最小发送间隔，取值范围为 100～1 000 整数毫秒。缺省值为 1 000ms
（2）min-rx-interval rx-value：可多选参数，指定 PIM BFD 报文的最小接收间隔，取值范围为 100～1 000 整数毫秒。缺省值为 1 000ms
（3）detect-multiplier multiplier-value：可多选参数，指定 PIM BFD 的本地检测倍数，取值范围为 3～50 的整数。缺省值为 3
缺省情况下，PIM BFD 报文的最小发送间隔、最小接收间隔都是 1 000ms；PIM BFD 的本地检测倍数为 3，可用 undo pim bfd { min-tx-interval tx-value | min-rx-interval rx-value | detect-multiplier multiplier-value } *命令恢复 PIM BFD 参数为缺省值</td></tr>
</table>

5．配置PIM GR

在堆叠系统中配置PIM GR（Graceful Restart）功能后，主交换机会向备交换机备份PIM路由表项、组播转发表以及需要向上游发送的Join/Prune信息。这样主备倒换后，新的主交换机就可以主动快速地向上游发送Join信息，维持上游的加入状态。具体配置步骤如表13-20所示。

表13-20 PIM GR的配置步骤

步骤	命令	说明
1	**system-view** 例如：<HUAWEI> **system-view**	进入系统视图
2	**pim** 例如：[HUAWEI] **pim**	键入要配置 BFD 的 PIM 接口，进入接口视图
3	**graceful-restart** 例如：[HUAWEI-pim] **graceful-restart**	全局使能 PIM GR。缺省情况下，没有使能 PIM GR 功能，**undo graceful-restart** 命令去使能 PIM GR 功能
4	**graceful-restart period** *period* 例如：[HUAWEI-pim] **graceful-restart period** 200	配置 PIM GR 的最小周期，用来保证转发过程中维持原有转发表项的最小时间，取值范围是 90～3 600 的整数秒 由于 PIM GR 是建立在单播 GR 的基础上的，因此配置 PIM GR 最小周期应大于所依赖单播 GR 的最小周期。**重复配置此命令将覆盖原有配置** 缺省情况下，PIM GR 最小周期为 120s，可用 **undo graceful-restart period** 命令恢复PIM GR最小周期的缺省值

13.3.6 PIM-SM管理

配置好PIM-SM（ASM或者SSM模型）后，可以通过一系列display任意视图命令PIM-SM相关功能及参数配置信息，如查看BSR、RP、PIM接口、PIM邻居和PIM路由表等信息，以验证配置及PIM-SM运行结果。

（1）使用display pimbsr-info命令查看BSR的信息。

（2）使用display pimrp-info [group-address] 命令查看所有或者指定组播组的RP信息。

（3）使用display pim interface [interface-type interface-number |up |down] [verbose] 命令查看所有或者指定接口，或者状态为Up或者Down接口上的摘要或者详细（选择verbose可选项时）PIM信息。

（4）使用 display pimneighbor [neighbor-address | interface interface-type interface- number | verbose] *命令查看指定邻居或者（和）接口上的PIM邻居信息。

（5）使用display pimrouting-table [group-address [mask {group-mask-length |group- mask }] | source-address [mask {source-mask-length |source-mask }] | incoming-interface {interface-typeinterface-number |register } |outgoing-interface {include |exclude |match } {interface- type interface-number |register |none } |mode {dm |sm |ssm } |flags flag- value |fsm] * [outgoing- interface-number [number]] 命令查看符合条件的PIM路由表的详细信息。

（6）使用 display pimrouting-tablebrief [group-address [mask {group-mask-length |group-mask }] |source-address [mask {source-mask-length |source-mask }] | incoming- interface { interface-type interface-number |register }] *命令查看符合条件的PIM路由表摘要信息。

（7）使用display pimbfd session statistics命令查看PIM BFD会话统计信息；使用display pimbfdsession [interface interface-type interface-number |neighborneighbor-address] *命令查看指定接口上或者与指定邻居之间的PIM BFD会话信息。

（8）使用display pim claimed-route [source-address] 命令查看所有或者指定组播组中PIM协议使用的单播路由信息。

（9）使用 display pim control-message countersmessage-type {probe | register |register-stop | crp }或 display pim control-message counters [message-type { assert | graft |graft-ack |hello | join-prune | state-refresh | bsr } |

interface interface-type interface- number] *命令查看发送、接收和无效的PIM控制报文数目。

（10）使用 display pim invalid-packet [interface interface-type interface-number |message-type { assert |bsr |hello | join-prune |graft | graft-ack | state-refresh }] *命令查看设备接收到的无效PIM报文的统计信息。

（11）使用 reset pim control-message counters [interface interface-type interface- number] 命令清除PIM控制报文统计信息。

13.3.7 PIM-SM（ASM模型）配置示例

本示例拓扑结构如图13-8所示，是一个单域PIM-SM网络。现用户主机HostA、HostB希望能够接收到Source发送的组播数据。

1．基本配置思路分析

本示例中没有明确要求用户仅接收指定组播源发来的数据，所以可以通过 PIM-SM ASM模型来实现，使得加入同一组播组的所有用户主机能够接收任意源发往该组的组播数据。总体配置任务如下（主要为PIM-SM ASM的基本功能配置）。

（1）配置交换机各VLAN接口IP地址和单播路由协议。组播域内路由协议PIM依赖单播路由协议，单播路由正常是组播协议正常工作的基础。

（2）在所有提供组播服务的交换机上使能组播路由功能，是配置PIM-SM的前提。

（3）在交换机所有接口上使能PIM-SM功能，然后才能配置PIM-SM的其他功能。

（4）在与主机侧相连的交换机接口上使能IGMP。组播组成员能通过发送IGMP消息自由加入或者离开某个组播组。叶节点交换机通过IGMP协议来维护组成员关系列表。

如果用户主机侧需同时配置PIM-SM和IGMP，必须先使能PIM-SM，再使能IGMP。

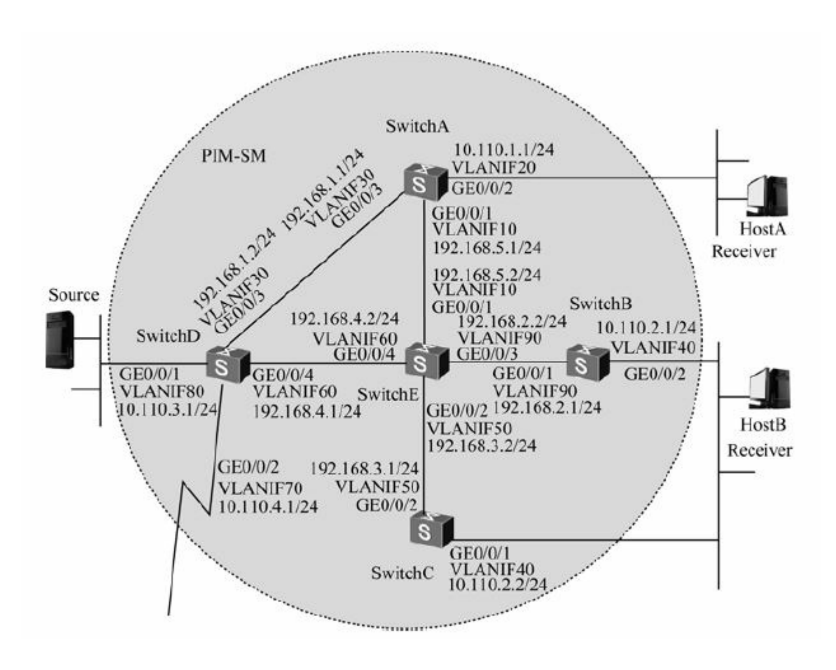


图13-8 ASM模型的PIM-SM域内组播配置示例

（5）可在与主机侧相连的交换机接口上使能PIM Silent，防止恶意主机模拟发送PIM Hello报文，增加PIM-SM域的安全性。但如果用户主机（如HostB）所在网段相连着多台交换机，那么这些交换机的用户主机侧接口不能使能PIM Silent，如图中的SwitchB、SwitchC的对应接口。

（6）配置RP。在PIM-SM域中，RP是提供ASM服务的核心，是转发组播数据的中转站。建议RP的位置配置在组播流量分支较多的交换机上，如图中的SwitchE的位置。

（7）在与外域相连的SwitchD GE0/0/1接口上配置BSR边界，自举报文不能通过该边界，使BSR只为该PIM-SM域服务，增加组播可控性。

2．具体配置步骤

下面是以上各配置任务的具体配置步骤。

（1）按照图中标注配置各交换机VLAN接口的IP地址和掩码，配置各交换机间采用OSPF进行互连，确保网络中各交换机间能够在网络层互通。因为SwitchA、SwitchB、SwitchC、SwitchD和SwitchE上的配置方法一样，所以下面仅以SwitchA上的配置为例进行介绍。

```
[SwitchA] vlan batch 10 20 30    #---批量创建VLAN 10、VLAN 20和VLAN 30
[SwitchA] interface gigabitethernet0/0/1
[SwitchA-GigabitEthernet0/0/1] port hybrid pvid vlan 10
[SwitchA-GigabitEthernet0/0/1] port hybrid untagged vlan 10
[SwitchA-GigabitEthernet0/0/1] quit
[SwitchA] interface gigabitethernet0/0/2
[SwitchA-GigabitEthernet0/0/2] port hybrid pvid vlan 20
[SwitchA-GigabitEthernet0/0/2] port hybrid untagged vlan 20
[SwitchA-GigabitEthernet0/0/2] quit
[SwitchA] interface gigabitethernet0/0/3
[SwitchA-GigabitEthernet0/0/3] port hybrid pvid vlan 30
[SwitchA-GigabitEthernet0/0/3] port hybrid untagged vlan 30
[SwitchA-GigabitEthernet0/0/3] quit
[SwitchA] interface vlanif 10
[SwitchA-Vlanif10] ip address 192.168.5.1 24
[SwitchA-Vlanif10] quit
[SwitchA] interfacevlanif 20
[SwitchA-Vlanif20] ip address 10.110.1.1 24
[SwitchA-Vlanif20] quit
[SwitchA] interfacevlanif 30
[SwitchA-Vlanif30] ip address 192.168.1.1 24
[SwitchA-Vlanif30] quit
[SwitchA] ospf
[SwitchA-ospf-1] area 0
[SwitchA-ospf-1-area-0.0.0.0] network10.110.1.0 0.0.0.255
[SwitchA-ospf-1-area-0.0.0.0] network192.168.1.0 0.0.0.255
[SwitchA-ospf-1-area-0.0.0.0] network192.168.5.0 0.0.0.255
[SwitchA-ospf-1-area-0.0.0.0] quit
[SwitchA-ospf-1] quit
```

（2）在所有交换机使能组播路由功能，在各VLAN接口上使能PIM-SM功能。同样因为SwitchA、SwitchB、SwitchC、SwitchD和SwitchE上的配置方法一样，所以也仅以SwitchA为例进行介绍。

[SwitchA] multicast routing-enable

[SwitchA] interfacevlanif 10

[SwitchA-Vlanif10] pim sm

[SwitchA-Vlanif10] quit

[SwitchA] interfacevlanif 20

[SwitchA-Vlanif20] pim sm

[SwitchA-Vlanif20] quit

[SwitchA] interfacevlanif 30

[SwitchA-Vlanif30] pim sm

[SwitchA-Vlanif30] quit

（3）在SwitchA连接用户主机的接口上使能IGMP功能。SwitchB和SwitchC上的配置过程与SwitchA上的配置相似，配置过程略。

[SwitchA] interfacevlanif 20

[SwitchA-Vlanif20] igmp enable

（4）在SwitchA接口上使能PIM silent。

[SwitchA] interfacevlanif 20

[SwitchA-Vlanif20] pim silent

（5）配置RP。配置RP有两种方式：静态RP和动态RP。可以同时配置，也可以只配置其中一种。同时配置两种RP时，可以通过参数调整优先选择哪种RP。本实例同时配置两种RP，缺省优选动态RP，静态RP作为备份。

下面是配置动态RP的方法，需要将PIM-SM域的一个或多个交换机上配置为C-RP和C-BSR。本例中指定SwitchE同时为C-RP和C-BSR，在SwitchE上配置RP服务的组地址范围，及C-BSR和C-RP所在接口位置。

[SwitchE] acl number2008

[SwitchE-acl-basic-2008] rule permit source225.1.1.0 0.0.0.255

[SwitchE-acl-basic-2008] quit

[SwitchE] pim

[SwitchE-pim] c-bsrvlanif 60

[SwitchE-pim] c-rp vlanif 60 group-policy 2008

下面是配置静态 RP 的方法，需要在所有交换机上指定静态 RP 的地址。因为SwitchA、SwitchB、SwitchC、SwitchD和SwitchE上的配置方法一样，下面仅以SwitchA上的配置为例进行介绍。

[SwitchA] pim

[SwitchA-pim] static-rp 192.168.2.2

（6）在SwitchD与外域相连的接口上配置BSR边界。

[SwitchD] interfacevlanif 70

[SwitchD-Vlanif70] pim bsr-boundary

[SwitchD-Vlanif70] quit

配置好后，可通过display pim interface命令查看接口上PIM的配置和运行情况，以验证配置结果。SwitchC上PIM的显示信息如下。

<SwitchC>display pim interface

VPN-Instance: public net

Interface	State	NbrCnt	HelloInt	DR-Pri	DR-Address
Vlanif40	up	0	30	1	10.110.2.2 (local)
Vlanif50	up	1	30	1	192.168.3.1 (local)

可通过display pim bsr-info命令查看交换机上BSR选举的信息。SwitchA和SwitchE上BSR信息分别如下（SwitchE上还显示C-BSR信息）。

```
<SwitchA>display pim bsr-info
VPN-Instance: public net
  Elected AdminScoped BSR Count: 0
  Elected BSR Address: 192.168.4.2
    Priority: 0
    Hash mask length: 30
    State: Accept Preferred
    Scope: Not scoped
    Uptime: 01:40:40
    Expires: 00:01:42
    C-RP Count: 1
<SwitchE>display pim bsr-info
VPN-Instance: public net
  Elected AdminScoped BSR Count: 0
  Elected BSR Address: 192.168.4.2
    Priority: 0
    Hash Mask length: 30
    State: Elected
    Scope: Not scoped
    Uptime: 00:00:18
    Next BSR message scheduled at :00:01:42
    C-RP Count: 1
  Candidate AdminScoped BSR Count: 0
  Candidate BSR Address: 192.168.4.2
    Priority: 0
    Hash mask length: 30
    State:Elected
    Scope: Not scoped
    Wait to be BSR: 0
```

可通过display pim rp-info命令查看Switch上获取的RP信息。SwitchA上RP信息如下。

```
<SwitchA>display pim rp-info
VPN-Instance: public net
  PIM-SM BSR RP Number:1
  Group/MaskLen: 225.1.1.0/24
    RP: 192.168.4.2
```

Priority: 0

Uptime: 00:45:13

Expires: 00:02:17

PIM SM static RP Number:1

Static RP: 192.168.2.2

可通过 display pim routing-table 命令查看 PIM 协议组播路由表。组播源（10.110.3.100/24）向组播组（225.1.1.1/24）发送信息，HostA、HostB 都加入了组播组（225.1.1.1/24）。SwitchA上的PIM组播路由表显示如下，其他交换机上组播路由表显示类似。

说明

缺省情况下，组成员端DR在收到组播源发来的第一份组播数据后就会触发SPT切换，新建（S，G）路由表项。因此交换机上显示的（S，G）路由表项一般都是 SPT切换后的（S，G）路由表项。

```
[SwitchA] display pim routing-table
VPN-Instance: public net
  Total 1 (*, G) entry; 1 (S, G) entry
(*, 225.1.1.1)
    RP: 192.168.4.2
  Protocol: pim-sm, Flag: WC
  UpTime: 00:13:46
  Upstream interface: Vlanif10,
    Upstream neighbor: 192.168.5.2
    RPF prime neighbor: 192.168.5.2
  Downstream interface(s) information:
  Total number of downstreams: 1
    1: Vlanif20
      Protocol: pim-sm, UpTime: 00:13:46, Expires:-
(10.110.3.100, 225.1.1.1)
    RP: 192.168.4.2
    Protocol: pim-sm, Flag: SPT ACT
    UpTime: 00:00:42
    Upstream interface: Vlanif30
      Upstream neighbor: 192.168.1.2
      RPF prime neighbor: 192.168.1.2
  Downstream interface(s) information:
    Total number of downstreams: 1
      1: Vlanif20
        Protocol: pim-sm, UpTime: 00:00:42, Expires:-
```

13.3.8 PIM-SM（SSM模型）配置示例

本示例拓扑结构如图13-9所示，是一个单域PIM-SM网络。现HostA希望能够接收组播源S1（10.110.4.100/24）、S2（10.110.3.100/24）发送的组播数据，而HostB希望能够接收组播源S2发送的组播

数据。

1．基本配置思路分析

本示例要求组播组成员仅接收指定源的组播数据，所以需要采用 PIM-SM SSM模型，使得用户主机在加入组播组的同时能够接收到自己所指定的组播源的组播数据。

相对于PIM-SM ASM模型来说，SSM模型的配置要简单许多，因为它既不需要维护RP，也不需要专门构建SPT，也无需注册组播源，仅需要使能SSM服务，配置用于限定接收指定组播源数据的组策略。下面是本示例的基本配置任务。

（1）配置交换机接口IP地址和单播路由协议。

（2）在所有提供组播服务的交换机上使能组播功能。

（3）在交换机所有接口上使能PIM-SM功能。

（4）在与主机侧相连的交换机接口上使能IGMP，并配置IGMP协议的版本号为v3，因为在不启用SSM Mapping的情况下仅 IGMPv3支持SSM服务。

如果用户主机侧需同时配置PIM-SM和IGMP，必须先使能PIM-SM，再使能IGMP。

（5）在与主机侧相连的交换机接口上使能 PIM Silent，防止恶意主机模拟发送 PIMHello报文，增加PIM-SM域的安全性。同样，如果用户主机所在网段相连着多台交换机，那么这些交换机的用户主机侧接口不能使能 PIM Silent，如图中的 SwitchB、SwitchC。

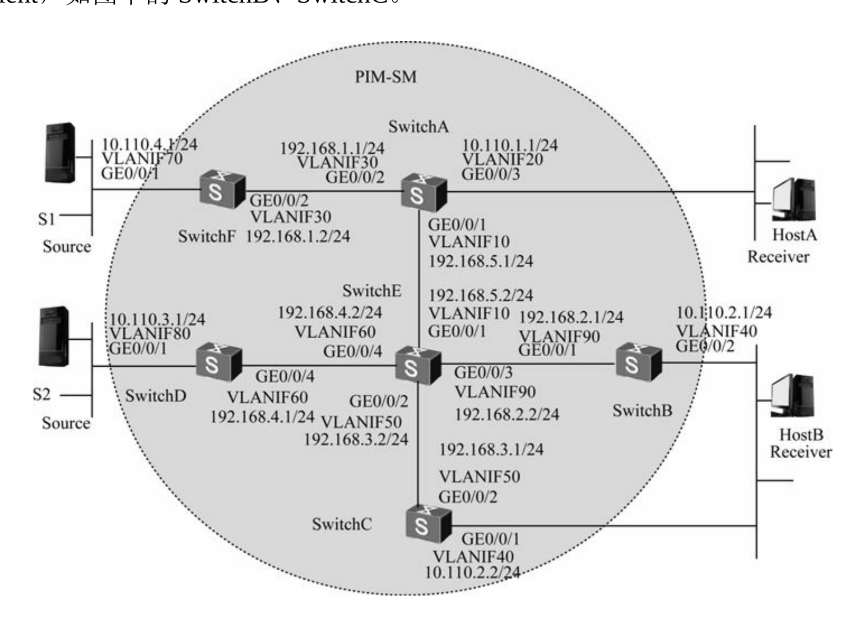


图13-9 SSM模型的PIM-SM域内组播配置示例拓扑结构

（6）在各交换机上设置 SSM 组地址范围。使 PIM-SM 域内的交换机为特定组地址范围内的SSM服务，实现可控组播。但各交换机上设置SSM组地址范围必须相同。

（7）在HostA和HostB主机连接的交换机VLANIF接口上配置Join-Prune报文过滤，以实现仅接收来自限定组播源的组播数据。

2．具体配置步骤

下面是本示例的具体配置步骤。

（1）按照图13-9中的标注配置各交换机VLAN接口的IP地址和掩码，配置各交换机间采用 OSPF 进行互连，确保网络中各交换机间能够在网络层互通。因为 SwitchA、SwitchB、SwitchC、SwitchD、SwitchE 和

SwitchF 上的配置方法一样，所以下面仅以SwitchA为例进行介绍。

```
[SwitchA] vlan batch 10 20 30
[SwitchA] interface gigabitethernet0/0/1
[SwitchA-GigabitEthernet0/0/1] port hybrid pvid vlan 10
[SwitchA-GigabitEthernet0/0/1] port hybrid untagged vlan 10
[SwitchA-GigabitEthernet0/0/1] quit
[SwitchA] interface gigabitethernet0/0/2
[SwitchA-GigabitEthernet0/0/2] port hybrid pvid vlan 20
[SwitchA-GigabitEthernet0/0/2] port hybrid untagged vlan 20
[SwitchA-GigabitEthernet0/0/2] quit
[SwitchA] interface gigabitethernet0/0/3
[SwitchA-GigabitEthernet0/0/3] port hybrid pvid vlan 30
[SwitchA-GigabitEthernet0/0/3] port hybrid untagged vlan 30
[SwitchA-GigabitEthernet0/0/3] quit
[SwitchA] interfacevlanif 10
[SwitchA-Vlanif10] ip address 192.168.5.1 24
[SwitchA-Vlanif10] quit
[SwitchA] interfacevlanif 20
[SwitchA-Vlanif20] ip address 10.110.1.1 24
[SwitchA-Vlanif20] quit
[SwitchA] interfacevlanif 30
[SwitchA-Vlanif30] ip address 192.168.1.1 24
[SwitchA-Vlanif30] quit
[SwitchA] ospf
[SwitchA-ospf-1] area 0
[SwitchA-ospf-1-area-0.0.0.0] network10.110.1.0 0.0.0.255
[SwitchA-ospf-1-area-0.0.0.0] network192.168.1.0 0.0.0.255
[SwitchA-ospf-1-area-0.0.0.0] network 192.168.5.0 0.0.0.255
[SwitchA-ospf-1-area-0.0.0.0] quit
[SwitchA-ospf-1] quit
```

（2）在所有交换机使能组播路由功能，在各VLAN接口上使能PIM-SM功能。同样因为 SwitchA、SwitchB、SwitchC、SwitchD、SwitchE和 SwitchF 的配置方法一样，所以下面也仅以SwitchA为例进行介绍。

```
[SwitchA] multicast routing-enable
[SwitchA] interfacevlanif 10
[SwitchA-Vlanif10] pim sm
[SwitchA-Vlanif10] quit
[SwitchA] interfacevlanif 20
[SwitchA-Vlanif20] pim sm
[SwitchA-Vlanif20] quit
```

[SwitchA] interface vlanif 30

[SwitchA-Vlanif30] pim sm

[SwitchA-Vlanif30] quit

（3）在SwitchA连接用户主机的接口上使能IGMPv3功能。SwitchB和SwitchC上的配置过程与SwitchA上的配置相似，配置过程略。

[SwitchA] interfacevlanif 20

[SwitchA-Vlanif20] igmp enable

[SwitchA-Vlanif20] igmp version 3

（4）在SwitchA接口上使能PIM silent。

[SwitchA] interfacevlanif 20

[SwitchA-Vlanif20] pim silent

（5）在所有交换机配置SSM组播组地址范围为232.1.1.0/24。因为SwitchA、SwitchB、SwitchC、SwitchD、SwitchE和SwitchF的配置方法一样，所以下面仅以SwitchA为例进行介绍。

[SwitchA] acl number2000

[SwitchA-acl-basic-2000] rule permit source232.1.1.0 0.0.0.255 #---限定232.1.1.0/24范围的组播组报文通过

[SwitchA-acl-basic-2000] quit

[SwitchA] pim

[SwitchA-pim] ssm-policy 2000

（6）在SwitchA的VLANIF20接口上配置Join-Prune报文过滤，指定HostA可接收组播源S1和S2发来的组播数据。

[SwitchA-pim] quit

[SwitchA] acl number 3001

[SwitchA-acl-adv-3001] rule permit source 10.110.3.100 0destination 232.1.1.0 0.0.255.255

[SwitchA-acl-adv-3001] rule permit source 10.110.4.100 0destination 232.1.1.0 0.0.255.255

[SwitchA-acl-adv-3001] quit

[SwitchA] interface vlanif 20

[SwitchA –Vlanif20] pim join-policy asm 3001

（7）在SwitchB和SwitchC的VLANIF40接口上配置Join-Prune报文过滤，指定HostB仅可接收组播源S2发来的组播数据。因SwitchB和SwitchC的一样，在此仅以SwitchB上的配置为例进行介绍。

[SwitchB] acl number 3001

[SwitchB-acl-adv-3001] rule permit source 10.110.3.100 0destination 232.1.1.0 0.0.255.255

[SwitchB-acl-adv-3001] quit

[SwitchB] interface vlanif 40

[SwitchB-Vlanif40] pim join-policy asm 3001

配置好后，可通过display pim interface命令查看接口上PIM的配置和运行情况，以验证配置结果。SwitchC上PIM的显示信息如下。

<SwitchC>display pim interface

VPN-Instance: public net

Interface State NbrCnt HelloInt DR-Pri DR-Address

Vlanif40 up 0 30 1 10.110.2.2 (local)

Vlanif50 up 1 30 1 192.168.3.1 (local)

可通过display pim routing-table命令查看PIM协议组播路由表。SwitchA和SwitchB上的显示信息如下。从中可以看出 HostA 接收了组播源（10.110.3.100/24）和组播源（ 10.110.4.100/24 ）发往组播组（ 232.1.1.1/24 ）的信息， HostB 只接收了组播源（10.110.3.100/24）发往组播组（232.1.1.1/24）的信息，达到了要求。

```
[SwitchA] display pim routing-table
  VPN-Instance: public net
  Total 2 (S, G) entry
(10.110.3.100, 232.1.1.1)
    Protocol: pim-ssm, Flag: SG_RCVCR
    UpTime: 00:13:46
    Upstream interface: Vlanif10,
      Upstream neighbor: 192.168.5.2
      RPF prime neighbor: 192.168.5.2
    Downstream interface(s) information:
    Total number of downstreams: 1
      1: Vlanif20
        Protocol: pim-ssm, UpTime: 00:13:46, Expires:-
(10.110.4.100, 232.1.1.1)
    Protocol: pim-ssm, Flag: SG_RCVCR
    UpTime: 00:00:42
    Upstream interface: Vlanif30
      Upstream neighbor: 192.168.1.2
      RPF prime neighbor: 192.168.1.2
  Downstream interface(s) information:
    Total number of downstreams: 1
      1: Vlanif20
        Protocol: pim-ssm, UpTime: 00:00:42, Expires:-
[SwitchB] display pim routing-table
  VPN-Instance: public net
  Total 1 (S, G) entry
(10.110.3.100, 232.1.1.1)
    Protocol: pim-ssm, Flag: SG_RCVCR
    UpTime: 00:10:12
    Upstream interface: Vlanif90,
      Upstream neighbor: 192.168.2.2
      RPF prime neighbor: 192.168.2.2
    Downstream interface(s) information:
```

Total number of downstreams: 1
  1: Vlanif40
    Protocol: pim-ssm, UpTime: 00:10:12, Expires:-

13.4 IGMP Snooping配置与管理

IGMP Snooping是一种 IPv4二层组播协议，通过侦听三层组播设备和用户主机之间发送的组播协议报文来维护组播报文的出端口信息，从而管理和控制组播数据报文在数据链路层的转发。

13.4.1 IGMP Snooping特性的产品支持

华为S系列交换机支持的 IGMP Snooping特性包括 IGMP Snooping基本功能、IGMP Snooping Proxy功能、IGMP Snooping策略、成员关系快速刷新以及 IGMP Snooping SSM Mapping等。IGMP Snooping作为一个二层组播特性，本章中涉及接口的配置，都是在二层物理接口（包括Eth-Trunk接口）下进行配置。仅**IGMP Snooping**基本功能必须配置，其他均为可选配置 。

1．IGMP Snooping基本功能

华为 S系列交换机所支持的基于VLAN的 IGMP Snooping（还有一种基于VSI的IGMP Snooping，本书不作介绍）的基本功能如下。

（1）支持IGMPv1、IGMPv2和IGMPv3，版本可配置。由于不同版本的IGMP协议报文不相同，因此需要为交换机配置和上游三层设备相同的版本。

（2）支持配置静态路由器端口和成员端口，实现组播数据快速稳定转发。

（3）支持配置 IGMP Snooping查询器功能，当上游没有启用 IGMP查询器时，交换机可以代替上游设备发送IGMP查询报文。

（4）支持 IGMP Snooping报文抑制功能，对成员主机上送的 IGMP Report和Leave报文进行抑制，从而降低上游设备的报文交互数量，提升系统性能。

（5）支持配置Router-Alert选项，提高设备性能以及网络安全性。

（6）支持配置抑制动态加入，禁止VLAN内收到的Report和Leave报文向配置有静态组的上游三层设备转发。

2．IGMP Snooping Proxy功能

通过在二层设备上配置 IGMP Snooping Proxy功能，可以同时具备报文抑制功能和查询器功能。配置了IGMP Snooping Proxy功能的交换机，在其上游设备看来，相当于一台主机；而在其下游主机看来，则相当于一台查询器。

3．IGMP Snooping策略

根据不同的场景要求，可以在交换机上进行如下配置，对组播报文进行过滤。

（1）通过配置组播组过滤策略，可以限制用户加入的组播组范围。

（2）通过配置接口可以学习的最大组播转发表项数量，可以控制接口上的组播数据流量。

（3）通过配置接口下组播数据过滤，可以拒绝从指定VLAN收到的组播数据。

（4）通过配置丢弃未知组播报文，使未知组播报文不在VLAN内广播。

4．成员关系快速刷新

成员关系快速刷新，即成员加入或者离开组播组时交换机快速响应成员变化，可以提高组播业务运行效率和用户体验。主要包括以下几个功能。

（1）调整动态成员端口老化时间。

（2）调整动态路由器端口老化时间。

（3）成员端口快速离开。

（4）二层网络拓扑变化时发送查询报文。

5．IGMP Snooping SSM Mapping

SSM提供了一种能够在成员端指定组播源的传输服务，需要IGMPv3的支持。如果某些组播组成员主机只能运行 IGMPv1或 IGMPv2，可以在交换机上配置 IGMP Snooping SSM Mapping功能，使组播组与组播源之间能够建立一一对应的映射关系，将 IGMPv1或IGMPv2报文中所包含的（*，G）信息映射为（S，G）信息，提供SSM组播服务。

以上 IGMP Snooping功能和参数的缺省配置如表 13-21所示。

表13-21 IGMP Snooping缺省配置

功能或参数	缺省值
IGMP Snooping 功能	未使能
IGMP Snooping 版本	IGMP Snooping 使能后，缺省的版本为 IGMPv2
IGMP Snooping 端口学习功能	IGMP Snooping 使能后，该功能缺省使能
IGMP Snooping 查询器	未使能
IGMP Snooping 报文抑制	未使能
IGMP Snooping Proxy	未使能
二层组播 SSM Mapping	未使能

13.4.2 IGMP Snooping基本功能配置任务

配置 IGMP Snooping基本功能，设备可以建立并维护二层组播转发表，实现组播数据报文在数据链路层的按需分发。在配置 IGMP Snooping基本功能之前，需完成连接接口并配置接口的物理参数，使其物理层状态为Up；创建VLAN并将对应接口加入VLAN中。具体包括以下配置任务。

1．使能 IGMP Snooping功能

使能全局 IGMP Snooping功能，是进行其他 IGMP Snooping配置的前提。VLAN下使能 IGMP Snooping功能，是VLAN下其他 IGMP Snooping配置生效的前提。

缺省情况下，交换机的全局 IGMP Snooping功能未使能。

2．配置 IGMP Snooping版本

IGMP 协议用于组成员关系管理，运行于三层组播设备和成员主机之间的网段。在二层设备上配置IGMP Snooping版本，设备可以处理相应版本的 IGMP报文。一般二层 设备上配置和三层组播设备一致的版本。如果三层组播设备没有启用 IGMP，则在二层设备上配置和成员主机相同或高于成员主机的版本。同一VLAN内必须运行同一个版本的 IGMP协议。如果VLAN内存在支持不同版本的主机，需要配置 IGMP Snooping版本，使设备可以处理所有主机的报文。

3．（可选）配置静态路由器端口

路由器端口是二层设备上朝向上游三层组播设备（组播路由器或三层交换机）的接口。VLAN内使能 IGMP Snooping功能后，加入该VLAN的接口会从组播协议报文中学习表项。当一个接口接收到 IGMP Query报文或PIM Hello报文时，二层设备会标识该接口为动态路由器端口。路由器端口主要有两个功能：一是接收上游的组播数据，二是指导 IGMP Report/Leave报文转发。当VLAN内收到 IGMP Report/Leave报文后，仅会向该VLAN内的路由器端口转发。

动态路由器端口会定时老化，当动态路由器端口在其老化时间超时前没有收到IGMP Query或者PIM

Hello报文，设备将把该接口从路由器端口列表中删除。如果希望某接口长期稳定地转发 IGMP Report/Leave报文到上游 IGMP查询器，可配置该接口为静态路由器端口。

4．（可选）配置静态成员端口

成员端口是设备上朝向组播组成员主机的接口，表示该接口下有组播组成员，可以通过组播协议动态学习或静态配置。VLAN内使能 IGMP Snooping功能后，加入该VLAN的接口会从组播协议报文中学习表项。当一个接口收到 IGMP Report报文时，设备会标识该接口为动态成员端口。动态成员端口会定时老化。

如果接口所连接的主机需要固定接收发往某组播组或组播源组的数据，可以配置该接口静态加入该组播组或组播源组，成为静态成员端口。静态成员端口不会老化。

5．（可选）配置 IGMP Snooping查询器

通过使能 IGMP Snooping，二层设备就可以通过侦听 IGMP查询器与用户主机间的IGMP 协议报文，动态建立二层组播转发表项，实现二层组播。但是当出现下面的情况时，即使二层设备运行了 IGMP Snooping，也会由于侦听不到 IGMP协议报文，而无法正常动态建立二层组播转发表项。

（1）上游三层组播设备在接口上未运行IGMP协议，而是配置了静态组播组。

（2）组播源和用户主机同属于一个二层网络，不需要三层组播设备。

此时，可通过在二层组播设备上配置 IGMP Snooping查询器，代替三层组播设备向用户主机发送 IGMP Query报文，从而解决此问题。

6.（可选）配置Report和Leave报文抑制

IGMP 协议通过周期性地查询和响应来维护组成员关系。在此过程中，如果多个成员加入了相同的组播组，会不断上送相同的Report报文给IGMP路由器。同时，当IGMPv2或IGMPv3的主机在离开某个组播组时，也会重复发送Leave报文。为了节约带宽，可以在二层设备上配置Report和Leave报文抑制功能。

当配置了对 Report 和 Leave 报文抑制后，针对每一个组播组，交换机会在第一次有成员加入需要建立组播表项，以及响应 IGMP 查询报文时，向上游转发一份 Report 报文；在最后一个组成员离开需要删除组播表项时，向上游转发一份 Leave报文。

7．（可选）配置Router-Alert选项

出于兼容性考虑，缺省情况下交换机不对Router-Alert选项进行检查，当收到IGMP报文时，不管其IP报头中是否携带Router-Alert选项，设备都会将其送给上层协议进行处理。为了提高系统性能、减少不必要的开支，同时出于协议安全性的考虑，可以配置对Router-Alert选项进行检查，当收到的IGMP报文中没有携带Router-Alert选项时，就丢弃该报文。

缺省情况下，交换机在发送的IGMP报文中携带Router-Alert选项。

8．（可选）配置 IGMP Snooping抑制动态加入

当上游三层设备为其他厂商设备，并且在用户主机侧接口上配置了静态组播组，不允许下游用户主机动态加入或离开组播组时，可以在设备上配置 IGMP Snooping抑制动态加入，禁止设备转发包含静态组地址信息的Report和Leave报文。

13.4.3 配置 IGMP Snooping基本功能

上节介绍的 IGMP Snooping基本功能的八项配置任务（只有前两项为必选的）的具体配置步骤如表13-22所示。

表13-22 IGMP Snooping基本功能配置步骤

配置任务	步骤	命令	说明
公共配置	1	**system-view** 例如：<HUAWEI> **system-view**	进入系统视图
使能IGMP Snooping功能	2	**igmp-snooping enable** 例如：[HUAWEI] **igmp-snooping enable**	使能全局 IGMP Snooping 功能 缺省情况下，全局 IGMP Snooping 功能均未使能，可用 **undo igmp-snooping enable** 命令禁止全局 IGMP Snooping 功能。如果禁止了全局 IGMP Snooping 功能，设备上所有 IGMP Snooping 相关配置将被删除。再次执行本命令使能全局 IGMP Snooping 功能后，设备上所有 IGMP Snooping 相关配置将被恢复为缺省配置
	3	**vlan** *vlan-id* 例如：[HUAWEI] **vlan** 10	键入要使能 IGMP Snooping 功能的 VLAN，进入 VLAN 视图
	4	**l2-multicast forwarding-mode { ip | mac }** 例如：[HUAWEI-vlan10] **l2-multicast forwarding-mode ip**	配置 VLAN 中组播流是按 IP 地址（选择 **ip** 二选一选项时）还是 MAC 地址（选择 **mac** 二选一选项时）转发 缺省情况下，S2700/5700S-LI/5700LI 系列交换机按 MAC 模式转发组播数据，其他 S 系列按 IP 模式转发组播数据，可用 **undo l2-multicast forwarding-mode** 命令恢复缺省情况

（续表）

配置任务	步骤	命令	说明
使能 IGMP Snooping 功能	5	**igmp-snooping enable** 例如：[HUAWEI-vlan10] **igmp-snooping enable**	使能 VLAN 的 IGMP Snooping 功能。使能了 VLAN 内 IGMP Snooping 之后，**该功能只会在已加入该 VLAN 的接口上生效，所以需要先把相应接口加入到此 VLAN 中** 可以在系统视图下使用 **igmp-snooping enable** [**vlan** { *vlan-id1* [**to** *vlan-id2*]} <1-10>]命令，使能多个 VLAN 的 IGMP Snooping 功能 IGMP Snooping 功能不能和 *N*：1（*N* 大于 1）VLAN Mapping 功能、VLAN Stacking 功能配合使用 缺省情况下，VLAN 的 IGMP Snooping 功能未使能，可用 **undo igmp-snooping enable** 命令去使能对应 VLAN 的 IGMP Snooping 功能
配置 IGMP Snooping 版本	6	**igmp-snooping version** *version* 例如：[HUAWEI-vlan10] **igmp-snooping version** 2	配置对应 VLAN 中的 IGMP Snooping 可以处理的 IGMP 版本，取值范围为 1～3 的整数。一般二层设备上配置和三层组播设备一致的版本。如果三层组播设备没有启用 IGMP，则在二层设备上配置和成员主机相同或高于成员主机的版本。当 VLAN 内存在支持不同版本的主机时，需执行本命令进行配置，使设备可以处理所有主机的报文 缺省情况下，设备可以处理 IGMPv1 和 IGMPv2 的报文，但无法处理 IGMPv3 的报文 【说明】当前面第 4 步配置的 VLAN 内的转发模式为基于 MAC 地址转发时，无法配置 IGMPv3 版本
（可选）配置静态路由器端口	7	**undo igmp-snooping router-learning** 例如：[HUAWEI-vlan10]**undo igmp-snooping router-learning**	（可选）禁止动态学习路由器端口，也可在对应物理接口视图下通过 **undo igmp-snooping router-learning vlan** { { *vlan-id1* [**to** *vlan-id2*] } <1-10> | **all** }命令禁止多个 VLAN 的动态路由器端口学习功能 缺省情况下，路由器端口动态学习功能处于使能状态，可用 **igmp-snooping router-learning** 命令使能 VLAN 的路由器端口动态学习功能
	8	**quit** 例如：[HUAWEI-vlan10] **quit**	退出 VLAN 视图，返回系统视图

（续表）

配置任务	步骤	命令	说明
(可选)配置静态路由器端口	9	**interface** *interface-type interface-number* 例如：[HUAWEI]**interface gigabitethernet 0/0/1**	键入要配置为静态路由器端口的**物理接口**，进入接口视图
	10	**igmp-snooping static-router-port vlan** { *vlan-id1* [**to** *vlan-id2*] } &<1-10> 例如：[HUAWEI-GigabitEthernet0/0/1] **igmp-snooping static-router-port vlan** 10	配置以上物理接口作为指定 VLAN 的静态路由器端口，命令中的参数说明如下。 （1）*vlan-id1* [**to** *vlan-id2*]：指定以上接口要作为单个 VLAN 或者一个范围（选择 **to** *vlan-id2* 可选参数时）的 VLAN 的路由器端口，VLAN ID 号的取值范围均为 1～4 094 （2）&<1-10>：表示 *vlan-id1* [**to** *vlan-id2*]参数对最多可以有 10 个 缺省情况下，接口没有配置为静态路由器端口，可用 **undo igmp-snooping static-router-port** 命令取消接口作为指定 VLAN 内的静态路由器端口
(可选)配置静态成员端口	11	**undo igmp-snooping learning vlan** { { *vlan-id1* [**to** *vlan-id2*] } &<1-10> | **all** } 例如：[HUAWEI-GigabitEthernet0/0/1] **undo igmp-snooping learning vlan** 10	禁止以上物理接口动态学习组播成员端口，命令中的参数说明如下。 （1）*vlan-id1* [**to** *vlan-id2*]：指定以上接口要禁止动态学习单个 VLAN 或者一个范围（选择 **to** *vlan-id2* 可选参数时）的 VLAN 中的组播成员端口，VLAN ID 号的取值范围均为 1～4 094 （2）&<1-10>：表示 *vlan-id1* [**to** *vlan-id2*]参数对最多可以有 10 个 （3）**all**：二选一选项，指定要禁止以上接口动态学习所有 VLAN 中的组播成员端口 禁止动态学习组播成员端口功能之后，如果要完成组播数据的转发，接口只能静态加入组播组 缺省情况下，成员端口动态学习功能处于使能状态，可用 **igmp-snooping learning** 命令恢复使能动态成员端口学习功能
	12	**l2-multicast static-group** [**source-address** *source-ip-address*] **group-address** *group-ip-address* **vlan** { *vlan-id1* [**to** *vlan-id2*] } &<1-10> 例如：[HUAWEI-GigabitEthernet0/0/1] **l2-multicast static-group -address** 224.1.1.1 **vlan** 10	配置以上物理接口静态加入对应组播组，成为对应组播组的静态成员端口。命令中的参数说明如下。 （1）*source-ip-address*：可选参数，指定要加入的组播组中的组播源 IP 地址，为单播 IP 地址 （2）*group-ip-address*：指定要加入的组播组 IP 地址，取值范围是 224.0.1.0～239.255.255.255 （3）*vlan-id1* [**to** *vlan-id2*] }：指定要静态加入组播组的 VLAN 范围 （4）&<1-10>：表示 *vlan-id1* [**to** *vlan-id2*] }参数最多有 10 个 也可以通过 **l2-multicast static-group** [**source-address** *source-ip-address*] **group-address** *group-ip-address1* **to** *group-ip-address2* **vlan** *vlan-id* 命令将接口批量加入多个组播组 缺省情况下，接口没有静态加入任何组播组，可用 **undo l2-multicast static-group** [**source-address** *source-ip-address*] **group-address** *group-ip-address* **vlan** { **all** | { *vlan-id1* [**to** *vlan-id2*] } &<1-10> }命令取消接口静态加入对应组播组的配置

（续表）

配置任务	步骤	命令	说明
（可选）配置 IGMP Snooping 查询器	13	**quit** 例如：[HUAWEI-GigabitEthernet0/0/1] **quit**	退出接口视图，返回系统视图
	14	**vlan** *vlan-id* 例如：[HUAWEI] **vlan** 10	键入要配置 IGMP Snooping 查询器的 VLAN，进入 VLAN 视图
	15	**igmp-snooping querier enable** 例如：[HUAWEI-vlan10] **igmp-snooping querier enable**	在以上 VLAN 中使能 IGMP Snooping 查询器功能。使能 IGMP Snooping 查询器功能后，交换机会定时以广播的方式向 VLAN 内所有接口（包括路由器端口）发送 IGMP Query 报文，如果组播网络中已经存在 IGMP 查询器，可能会引起 IGMP 查询器重新选举。此时，建议不配置此功能。 【注意】如果与 VLAN 对应的三层 VLANIF 接口使能了 IGMP 功能，则不能在该 VLAN 内使能 IGMP Snooping 查询器功能 另外，在同一 VLAN 内，IGMP Snooping 查询器功能和 IGMP Snooping Proxy 功能不能同时配置。如果设备上配置了组播 VLAN 复制功能，也不能在用户 VLAN 上使能 IGMP Snooping 查询器功能 缺省情况下，VLAN 内没有使能 IGMP Snooping 查询器功能，可用 **undo igmp-snooping querier enable** 命令去使能对应 VLAN 的 IGMP Snooping 查询器功能
	16	**igmp-snooping query-interval** *query-interval* 例如：[HUAWEI-vlan10] **igmp-snooping query-interval** 300	（可选）配置 VLAN 内的 IGMP 普遍组查询报文发送时间间隔，取值范围为 1～65 535 的整数秒 缺省情况下，VLAN 内的 IGMP 普遍查询报文发送时间间隔为 125s，可用 **undo igmp-snooping query-interval** 命令恢复 VLAN 内的 IGMP 普遍组查询报文发送时间间隔为缺省值
	17	**igmp-snooping robust-count** *robust-count* 例如：[HUAWEI-vlan10] **igmp-snooping robust-count** 3	（可选）配置 VLAN 内的 IGMP 健壮系数，即发送 Query 报文的次数，取值范围为 2～5 的整数。健壮系数用来规定以下两个值。 （1）当查询器启动时发送“健壮系数”次的“普遍组查询报文”，发送时间间隔为“普遍组查询报文发送间隔”的 1/4 （2）当设备收到 Leave 报文后，发送“健壮系数”次的“IGMP 特定组查询报文”，发送间隔为“特定组查询报文发送间隔” 缺省情况下，VLAN 内的 IGMP 健壮系数为 2，可用 **undo igmp-snooping robust-count** 命令恢复 VLAN 内的 IGMP 健壮系数为缺省值

（续表）

配置任务	步骤	命令	说明
(可选)配置 IGMP Snooping 查询器	18	**igmp-snooping max-response-time** *max-response-time* 例如：[HUAWEI-vlan10] **igmp-snooping max-response-time** 20	（可选）在 VLAN 内配置 IGMP 普遍组查询的最大响应时间，取值范围为 1～25 的整数秒。**但 IGMPv1 不支持** 【说明】配置本命令后，当交换机收到主机的 IGMP Report 报文后，成员端口老化时间设置为普遍组查询报文的发送间隔 × IGMP 健壮系数 + 最大响应时间；组播组成员接收到一个 IGMP 查询报文后，会在最大响应时间内发送 Report 报文 缺省情况下，VLAN 内的 IGMP 普遍组查询最大响应时间为 10s，可用 **undo igmp-snooping max-response-time** 命令恢复 VLAN 内 IGMP 普遍组查询的最大响应时间缺省值
	19	**igmp-snooping lastmember-queryinterval** *lastmember-queryinterval* 例如：[HUAWEI-vlan10] **igmp-snooping lastmember-queryinterval** 3	（可选）配置 VLAN 内的最后成员查询时间间隔，即 IGMP 特定组查询报文发送时间间隔，取值范围为 1～5 整数秒。**但 IGMPv1 不支持** 【说明】配置本命令后，当交换机收到主机退出某组播组的 Leave 报文时，重置成员端口老化时间为特定组查询报文发送间隔 × IGMP 健壮系数。即会连续发送"IGMP 健壮系数"次特定组成员查询报文，询问该组播组是否还存在成员。本参数定义了发送该报文的时间间隔 缺省情况下，VLAN 内的 IGMP 特定组查询报文发送时间间隔为 1s，可用 **undo igmp-snooping lastmember-queryinterval** 命令恢复 VLAN 内的最后成员查询时间间隔为缺省值
	20	**quit** 例如：[HUAWEI-vlan10] **quit**	（可选）退出 VLAN 视图，返回系统视图
	21	**igmp-snooping send-query source-address** *ip-address* 例如：[HUAWEI] **igmp-snooping send-query source-address** 1.1.1.1	（可选）配置 IGMP 普遍组查询报文的组播源 IP 地址 缺省情况下，IGMP Snooping 查询器发送普遍组查询报文时源 IP 地址为 192.168.0.1，当该地址已被网络中的其他设备占用时，可使用本命令配置为其他地址。可用 **undo igmp-snooping send-query source-address** 命令恢复 IGMP 普遍组查询报文的源 IP 地址为缺省值
(可选)配置 Report 和 Leave 报文抑制	22	**vlan** *vlan-id* 例如：[HUAWEI]**vlan** 10	键入要配置报文抑制功能的 VLAN，进入 VLAN 视图
	23	**igmp-snooping report-suppress** 例如：[HUAWEI-vlan10] **igmp-snooping report-suppress**	配置在 VLAN 内对 Report 和 Leave 报文的抑制功能 【说明】配置此功能需注意以下几点。 (1) 在某 VLAN 下配置了报文抑制功能后，不能在与之对应的三层 VLANIF 接口上使能 IGMP 功能 (2) 在同一 VLAN 内，Report 和 Leave 报文抑制功能和 IGMP Snooping Proxy 不能同时配置

（续表）

配置任务	步骤	命令	说明
（可选）配置 Report 和 Leave 报文抑制	23	**igmp-snooping report-suppress** 例如：[HUAWEI-vlan10] **igmp-snooping report-suppress**	（3）如果设备上配置了组播 VLAN 复制功能，则不能在用户 VLAN 上配置 Report 和 Leave 报文抑制功能 （4）设备未使能报文抑制功能时，对重复的成员关系报告报文也会进行抑制，缺省的抑制时间为 10s，此时间可通过 **igmp-snooping suppress-time** *suppress-time* 命令来配置。如果将 *suppress-time* 设为 0，表示对所有的成员关系报文都立即转发 缺省情况下，**undo igmp-snooping report-suppress** 命令取消在 VLAN 内对 Report 和 Leave 报文的抑制
（可选）配置 Router-Alert 选项	24	**igmp-snooping require-router-alert** 例如：[HUAWEI-vlan10] **igmp-snooping require-router-alert**	配置设备对接收的IGMP报文进行Router-Alert检查 【说明】Router-Alert 是一种标识协议报文的特殊机制，如果一个报文中带有 Router-Alert 选项，则表示该报文需要被上送到路由协议层去处理。出于兼容性考虑，缺省情况下设备不对 Router-Alert 选项进行检查，IGMP 报文中无论是否携带有 Router-Alert 选项，设备都会将其送给上层协议进行处理。为了提高设备性能、减少不必要的开支，同时出于协议安全性的考虑，可以配置设备丢弃未携带 Router-Alert 选项的 IGMP 报文，此时当设备收到IGMP报文时，会检查该报文的Router-Alert 选项，如果没有携带该选项，就丢弃该报文 可用 **undo igmp-snooping require-router-alert** 命令恢复缺省配置
	25	**igmp-snooping send-router-alert** 例如：[HUAWEI-pim]	配置设备发送的IGMP报文中携带Router-Alert选项。缺省情况下设备不对 Router-Alert 选项进行检查，可用 **undo igmp-snooping send-router-alert** 命令恢复缺省配置
（可选）配置 IGMP Snooping 抑制动态加入	26	**igmp-snooping static-group suppress-dynamic-join** 例如：[HUAWEI-vlan10] **igmp-snooping static-group suppress-dynamic-join**	禁止VLAN内收到的包含有静态组地址信息的 Report 和 Leave 报文向配置该静态组的上游三层设备转发 【说明】如果二层设备的上游三层组播设备为其他厂商设备，并且在此三层设备的接口上配置了静态组播组，不允许用户以动态的方式加入或者退出组播组，此时需要在二层设备上配置禁止向三层组播设备转发包含有静态组地址信息的 Report 和 Leave 报文 缺省情况下，VLAN 内收到的包含有静态组地址信息的 Report 和 Leave 报文向配置该静态组的上游三层设备转发，可用 **undo igmp-snooping static-group suppress-dynamic-join** 命令恢复缺省配置

【示例 1】配置VLAN2内的GE0/0/1接口静态加入组播组224.1.1.1。

```
<HUAWEI>system-view
[HUAWEI] igmp-snooping enable
[HUAWEI] interface gigabitethernet 0/0/1
[HUAWEI-GigabitEthernet0/0/1] port link-type trunk
[HUAWEI-GigabitEthernet0/0/1] port trunk allow-pass vlan 2
[HUAWEI-GigabitEthernet0/0/1] l2-multicast static-group group-address 224.1.1.1 vlan 2
```

【示例 2】配置VLAN2内的GE0/0/1接口加入组播组224.1.1.1～224.1.1.3。

```
<HUAWEI>system-view
[HUAWEI] igmp-snooping enable
[HUAWEI] interface gigabitethernet 0/0/1
[HUAWEI-GigabitEthernet0/0/1] port link-type trunk
[HUAWEI-GigabitEthernet0/0/1] port trunk allow-pass vlan 2
```

[HUAWEI-GigabitEthernet0/0/1] l2-multicast static-group group-address 224.1.1.1 to 224.1.1.3 vlan 2

【示例 3】取消GE0/0/1接口静态加入所有VLAN的组播组224.1.1.1。

<HUAWEI> system-view

[HUAWEI] interfacegigabitethernet 0/0/1

[HUAWEI-GigabitEthernet0/0/1] undo l2-multicast static-group group-address 224.1.1.1 vlan all

13.4.4 配置 IGMP Snooping Proxy

IGMP Snooping Proxy功能在 IGMP Snooping的基础上使交换机代替上游三层设备向下游主机发送IGMP Query报文和代替下游主机向上游设备发送 IGMP Report和Leave报文，这样能够有效地节约上游设备和本设备之间的带宽。

当三层设备没有启用IGMP时（如只配置了静态组播组），网络中就不会有IGMP查询器来维护组成员关系。通过在二层设备上配置 IGMP Snooping Proxy功能可以使其发送Query报文，充当IGMP查询器。当网络中运行了IGMP时，为了减少上游三层设备收到的 IGMP Report报文和Leave报文数量，也可在二层设备上部署 IGMP Snooping Proxy功能，使其能够代理下游主机来向上游设备发送成员关系报告报文。IGMP Snooping Proxy功能的具体配置步骤如表13-23所示。

表13-23 IGMP Snooping Proxy功能配置步骤

步骤	命令	说明
1	**system-view** 例如：<HUAWEI> **system-view**	进入系统视图
2	**vlan** *vlan-id* 例如：[HUAWEI]**vlan** 10	键入要配置 IGMP Snooping Proxy 功能的 VLAN，进入VLAN 视图
3	**igmp-snooping proxy** 例如：[HUAWEI-vlan10] **igmp-snooping proxy**	使能 IGMP Snooping Proxy 功能 缺省状况下，VLAN 内没有使能 IGMP Snooping Proxy 功能，可用 **undo igmp-snooping proxy** 命令去使能 VLAN 内的 IGMP Snooping Proxy 功能
4	**quit** 例如：[HUAWEI-vlan10] **quit**	退出接口视图，返回系统视图
5	**Interface** *interface-type interface-number* 例如：[HUAWEI]**interface gigabitethernet 0/0/1**	键入要配置 IGMP Snooping Proxy 功能的物理接口，进入接口视图

（续表）

步骤	命令	说明
6	**igmp-snooping proxy-uplink-port vlan** *vlan-id* 例如：[HUAWEI-GigabitEthernet0/0/1] **igmp-snooping proxy-uplink-port vlan** 10	配置 IGMP Snooping Proxy 上行接口，禁止向此接口发送IGMP 查询报文。参数 *vlan-id* 用来指定以上接口作为 IGMP Snooping Proxy 上行接口的 VLAN ID，取值范围为 1～4 094 启用 IGMP Snooping Proxy 功能后，交换机会定时以广播的方式向 VLAN 内所有接口（包括路由器端口）发送 IGMP Query 报文，可能会引起 IGMP 查询器重新选举。当上游已经启用 IGMP 时，配置此命令可以禁止交换机向路由器端口转发 Query 报文，避免查询器重新选举 缺省情况下，VLAN 没有配置 IGMP Snooping Proxy 上行接口，可用 **undo igmp-snooping proxy-uplink-port** 命令删除 IGMP Snooping Proxy 上行接口

13.4.5 配置 IGMP Snooping策略

通过配置 IGMP Snooping策略可以控制用户对组播节目的点播，提高二层组播网络的可控性和安全性，相当于本章前面13.1.3节介绍的IGMP过滤策略。本功能需要结合ACL使用，所以先创建ACL并在其规

则中定义组播组过滤策略。

组播组过滤策略主要用于对VLAN内的主机加入的组播组进行限制。本功能仅对动态加入的组生效，对静态组播组无效。具体的配置方法是在对应的 VLAN 视图下使用igmp-snooping group-policyacl-number [version version-number] [default-permit] 命令配置当前VLAN的组播组过滤策略。命令中的参数和选项说明如下。

（1）acl-number：指定用来定义该VLAN内用户主机可以加入的组播组范围的ACL的编号，可以是基本ACL（源地址就是组播组IP地址，表示仅过滤组播组IP地址），取值范围为 2 000～2 999，和高级ACL（源地址为组播源 IP地址，目的地址是组播组 IP地址，表示同时过滤组播源IP地址和组播组IP地址），取值范围为3000～3999。

（2）version-number：可选参数，指定IGMP报文的版本，表示只对指定版本的IGMP报文应用组播组过滤策略。如果不指定该参数，则设备对接收到的所有IGMP报文都应用该组播组过滤策略。

（3）default-permit：可选项，指定组播组过滤策略的缺省行为是对所有组播组许可，即表示如果引用的ACL未定义规则，则允许VLAN内用户主机可以加入所有组播组。如果组播组过滤策略未指定default-permit可选项，则rule命令必须使用permit选项允许VLAN内的主机访问指定组播组，完成过滤组播组的目的；如果组播组过滤策略指定了default-permit可选项，则rule命令必须使用deny选项明确禁止VLAN内的主机访问指定组播组，完成过滤组播组的目的。

缺省状况下，VLAN无组播组过滤策略，即VLAN内的用户主机可以加入任何组播组，可用undo igmp-snooping group-policy命令取消当前VLAN的组播组过滤策略。

【示例 1】禁止VLAN 2内的用户主机加入组播组 225.1.1.123。

```
<HUAWEI>system-view
[HUAWEI] acl number 2000
[HUAWEI-acl-basic-2000] rule deny source225.1.1.123 0
[HUAWEI-acl-basic-2000] quit
[HUAWEI] igmp-snooping enable
[HUAWEI] vlan 2
[HUAWEI-vlan2] igmp-snooping enable
[HUAWEI-vlan2] igmp-snooping group-policy 2000default-permit
```

【示例 2】允许VLAN 2的用户主机加入组播源组（10.10.10.1，225.1.1.123）。

```
<HUAWEI>system-view
[HUAWEI] acl number 3000
[HUAWEI-acl-adv-3000] rule permit source 10.10.10.1 225.1.1.123 0
[HUAWEI-acl-adv-3000] quit
[HUAWEI] igmp-snooping enable
[HUAWEI] vlan 2
[HUAWEI-vlan2] igmp-snooping enable
[HUAWEI-vlan2] igmp-snooping group-policy 3000
```

13.4.6 配置接口下组播数据过滤

当网络管理员希望拒绝某特定的组播数据报文时，可以在交换机接口下配置组播数据过滤，拒绝来自指定VLAN的组播数据报文。配置方法很简单，只需在对应的接口视图下使用multicast-source-deny vlan {

vlan-id1 [to vlan-id2] } &<1-10>命令使接口对指定VLAN（接口必须已加入到对应的VLAN中）内的组播数据进行过滤即可。命令中的参数说明如下。

（1）vlan-id1：指定要过滤组播报文的起始VLAN的VLAN ID，取值范围为1～4 094。

（2）to vlan-id2：可选参数，指定要过滤组播报文的结束VLAN的VLAN ID，取值范围也为1～4 094，但必须大于参数vlan-id1的值，与vlan-id1共同指定一个范围的VLAN。

（3）&<1-10>：表示前面的 vlan-id1 [to vlan-id2] 可以最多有 10个。

在接口下配置本命令后，接口会丢弃收到的指定VLAN的组播报文。在如下场景下，可能会需要使用此功能。

（1）用户侧接口上收到了组播报文，而交换机一般不需要接收来自用户侧接口的组播数据报文。在用户侧接口配置本命令，丢弃该接口收到的组播数据，可以防止用户主机恶意伪造组播源发送组播流。

（2）不同VLAN的多个组播源和交换机之间二层相连，交换机只想接收部分源的数据。

（3）在某些特殊情况下，比如某接口下用户组播业务到期需要暂时停止，网络管理员可以通过配置本命令，来实现拒绝相应VLAN的组播数据报文。

但使用此命令只过滤同时满足以下条件的组播数据报文。

（1）报文目的MAC地址为IP组播MAC地址（即0x01-00-5e开头的IPv4组播MAC地址或0x3333开头的IPv6组播MAC地址）。

（2）报文封装的协议类型为UDP类型。

【示例 】在GE1/0/1接口上配置对VLAN100到VLAN105的组播数据丢弃处理。

```
<HUAWEI>system-view
[HUAWEI] interface gigabitethernet1/0/1
[HUAWEI-GigabitEthernet1/0/1] multicast-source-deny vlan 100 to 105
```

13.4.7 配置丢弃未知组播流

所谓“未知组播流”就是组播转发表中不存在对应表项的组播报文。缺省情况下，交换机对未知组播流的处理方式为在VLAN内广播。通过配置丢弃未知组播流，可以节省瞬时带宽占用率。配置方法很简单，只需在对应的 VLAN 视图下使用 multicast drop-unknown 命令配置丢弃未知组播流。但要注意的是，配置本命令后会丢弃一切未知组播报文，包括在VLAN内透传的使用保留组播地址的协议报文。

13.4.8 配置成员关系快速刷新

配置成员关系快速刷新，使组播组成员加入或者离开组播组时设备能够快速响应成员变化，可以提高组播业务运行效率和用户体验。可以进行以下几方面的配置。

1．配置动态成员端口老化时间

设备在收到不同IGMP协议报文之后，会为成员端口启动不同时长的老化定时器。

（1）当设备的成员端口收到下游主机的Report报文后，将接口老化时间设置为健壮系数×普遍组查询报文发送时间间隔+最大响应时间。

（2）当设备的成员端口收到下游主机的Leave报文后，将接口老化时间设置为特定组查询报文发送时间间隔×健壮系数。

动态成员端口老化时间的具体配置步骤如表13-24所示。

表13-24 动态成员端口老化时间的配置步骤

步骤	命令	说明
1	**system-view** 例如：<HUAWEI> **system-view**	进入系统视图
2	**vlan** *vlan-id* 例如：[HUAWEI]**vlan** 10	键入要配置动态成员端口老化时间的 VLAN，进入 VLAN 视图
3	**igmp-snooping query-interval** *query-interval* 例如：[HUAWEI-vlan10] **igmp-snooping query-interval** 100	配置查询器发送普遍组查询报文的时间间隔，取值范围为 1～65 535 的整数秒 缺省情况下，普遍组查询时间间隔为 60s，可用 **undo igmp-snooping query-interval** 命令恢复对应 VLAN 内的 IGMP 普遍组查询报文发送时间间隔为缺省值
4	**igmp-snooping robust-count** *robust-count* 例如：[HUAWEI-vlan10] **igmp-snooping robust-count** 3	配置查询器的 IGMP 健壮系数，取值范围为 2～5 的整数 缺省情况下，IGMP 健壮系数为 2，可用 **undo igmp-snooping robust-count** 命令恢复对应 VLAN 内的 IGMP 健壮系数为缺省值
5	**igmp-snooping max-response-time** *max-response-time* 例如：[HUAWEI-vlan10] **igmp-snooping max-response-time** 20	配置查询器最大响应时间，取值范围为 1～25 的整数秒 缺省情况下，IGMP 查询报文的最大响应时间是 10s，可用 **undo igmp-snooping max-response-time** 命令恢复对应 VLAN 内 IGMP 普遍组查询的最大响应时间为缺省值
6	**igmp-snooping lastmember-queryinterval** *lastmember-queryinterval* 例如：[HUAWEI-vlan10] **igmp-snooping lastmember-queryinterval** 3	配置 VLAN 内的最后成员查询时间间隔，即 IGMP 特定组查询报文发送时间间隔，取值范围为 1～5 的整数秒 缺省情况下，特定组查询时间间隔为 1s，可用 **undo igmp-snooping lastmember-queryinterval** 命令恢复 VLAN 内的最后成员查询时间间隔为缺省值

2．配置动态路由器端口老化时间

IGMP Snooping的路由器端口用来向上游三层设备发送Report报文和接收上游设备的组播数据报文。在配置 IGMP Snooping功能后，设备可以动态学习路由器端口，实时监测上游组播数据的下发。当网络发生拥塞或者网络稳定性不佳时，动态路由器端口在其老化时间超时前没有收到 IGMP普遍组查询报文或者PIM Hello报文，设备将把该接口从路由器端口列表中删除，可能造成组播数据中断，此时可以将路由器端口老化时间值适当调大。

配置动态路由器端口老化时间的配置方法很简单，只需在对应的VLAN视图下使用igmp-snooping router-aging-time router-aging-time命令即可，取值范围为 1～1 000的整数秒。缺省情况下，通过IGMP普遍组查询报文学习到的路由器端口老化时间为180s；通过PIM Hello报文学习到的路由器端口老化时间为Hello报文中Holdtime值，可用undo igmp-snooping router-aging-time命令恢复对应VLAN内的动态路由器端口老化时间为缺省值。

【示例 1】配置VLAN3内的动态路由器端口老化时间为300s。

```
<HUAWEI> system-view
[HUAWEI] igmp-snooping enable
[HUAWEI] vlan 3
[HUAWEI-vlan3] igmp-snooping enable
[HUAWEI-vlan3] igmp-snooping router-aging-time 300
```

3．配置成员端口快速离开

成员端口快速离开是指当交换机从成员端口接收到 IGMP Leave报文时，不再启动老化定时器等待转发表项老化，而是立即将该接口对应的转发表项删除。

注意

只有当 VLAN 内的每个接口下都只有一个组播组成员主机时，才可以使能该VLAN的成员端口快速离开功能。只有当交换机在VLAN内可以处理IGMPv2或IGMPv3报文时，配置成员端口快速离开功能才有意义。

配置成员端口快速离开功能的方法也很简单，只需在对应的 VLAN 视图下使用igmp-snooping prompt-leave [group-policy acl-number [default-permit]] 命令配置允许VLAN内的成员端口快速离开组播组即可。命

令中的参数和选项说明如下。

（1）acl-number：可选参数，用来指定要允许端口快速离开某些组播组，可以是基本ACL（源地址为组播 IP地址），取值范围为 2 000～2 999，也可以是高级ACL（源地址是组播源 IP地址，目的地址为组播组 IP地址），取值范围为 3 000～3 999。不指定此参数时表示所有组播组都允许端口快速离开。

（2）default-permit：可选项，则端口快速离开组播组策略的缺省行为是允许端口快速离开所有组播组。此时需要在ACL的rule命令中使用deny选项明确禁止成员端口快速离开指定组播组。如果不指定此可选项，则端口快速离开组播组策略的缺省行为是禁止端口快速离开所有组播组。此时需要通过 rule 命令中的 permit 选项实现只允许成员端口快速离开指定组播组。

缺省情况下，不允许成员端口快速离开组播组，可用undo igmp-snooping prompt-leave命令禁止VLAN内的成员端口快速离开组播组。

【示例 2】配置允许VLAN3内的成员端口快速离开组播组225.1.1.123。

```
<HUAWEI>system-view
[HUAWEI] igmp-snooping enable
[HUAWEI] acl number 2000
[HUAWEI-acl-basic-2000] rule permit source 225.1.1.123 0
[HUAWEI-acl-basic-2000] quit
[HUAWEI] vlan 3
[HUAWEI-vlan3] igmp-snooping enable
[HUAWEI-vlan3] igmp-snooping prompt-leave group-policy 2000
```

【示例 3】配置禁止VLAN3内的成员端口快速离开组播组225.1.1.123。

```
<HUAWEI>system-view
[HUAWEI] igmp-snooping enable
[HUAWEI] acl number 2000
[HUAWEI-acl-basic-2000] rule deny source 225.1.1.123 0
[HUAWEI-acl-basic-2000] quit
[HUAWEI] vlan 3
[HUAWEI-vlan3] igmp-snooping enable
[HUAWEI-vlan3] igmp-snooping prompt-leave group-policy 2000 default-permit
```

4．配置网络拓扑变化时发送Query报文

当二层网络拓扑发生变化时，组播报文的转发路径可能发生变化。配置交换机在链路故障时主动发送 IGMP Query报文，当组播组成员回应 IGMP Report报文时，设备根据Report报文更新成员端口信息，将组播数据流迅速切换到新的转发路径上。

在网络拓扑变化时发送Query报文的配置步骤如表13-25所示。

表13-25 在网络拓扑变化时发送Query报文的配置步骤

步骤	命令	说明
1	**system-view** 例如：<HUAWEI> **system-view**	进入系统视图
2	**igmp-snooping send-query enable** 例如：[HUAWEI]**igmp-snooping send-query enable**	配置设备在网络拓扑变化时发送 IGMP 普遍组查询报文。配置本命令后，当设备感知二层网络拓扑发生变化时，会主动发送 IGMP 普遍组查询报文（报文源地址缺省为 192.168.0.1），保证设备能够快速更新端口信息，减少下游组成员接收组播数据中断时间 缺省情况下，当网络拓扑变化时，设备不会主动发送 IGMP 普遍组查询报文，可用 **undo igmp-snooping send-query enable** 命令禁止设备响应二层拓扑变化主动发送 IGMP 普遍组查询报文
3	**igmp-snooping send-query source-address** *ip-address* 例如：[HUAWEI] **igmp-snooping send-query source-address** 1.1.1.1	（可选）配置 IGMP 普遍组查询报文的源 IP 地址。 缺省情况下，响应拓扑变化时发送的普遍组查询报文源地址为 192.168.0.1。当该地址已被网络中的其他设备占用时，可使用本命令配置为其他地址，可用 **undo igmp-snooping send-query source-address** 命令恢复 IGMP 普遍组查询报文的源 IP 地址为缺省值

13.4.9 配置 IGMP Snooping SSM Mapping

在二层网络中，如果某些用户主机只能运行IGMPv1或IGMPv2，但是这些用户希望享受SSM服务，就需要在设备上配置 IGMP Snooping SSM Mapping功能。如果需要改变SSM组地址范围，则还可以配置SSM组策略，以使对应的组播组被作为SSM组播组对待。

1．（可选）配置SSM组策略

缺省情况下，SSM组范围是232.0.0.0～232.255.255.255。如果用户加入的组播组地址不在 SSM范围内，需要先在VLAN上配置SSM组策略，将组播组地址加入到SSM组地址范围。

SSM组策略的具体配置方法是在对应的VLAN视图下使用 igmp-snooping ssm-policy basic-acl-number命令，通过基本ACL（源地址为组播组IP地址）来限定允许作为SSM范围内的组对待的组播组。但此时ACL中的rule命令必须使用permit选项指定组播组IP地址才能生效；如果使用deny选项或指定的地址不是组播组IP地址，配置不生效。

【示例 1】配置VLAN3内组地址225.1.1.123作为SSM范围内的组。

```
<HUAWEI> system-view
[HUAWEI] acl number 2000
[HUAWEI-acl-basic-2000] rule permit source 225.1.1.123 0
[HUAWEI-acl-basic-2000] quit
[HUAWEI] igmp-snooping enable
[HUAWEI] vlan 3
[HUAWEI-vlan3] igmp-snooping enable
[HUAWEI-vlan3] igmp-snooping ssm-policy 2000
```

2．配置 IGMP Snooping SSM Mapping

使能IGMP Snooping SSM Mapping功能后可以使组播组与组播源之间建立一一对应的映射关系。配置VLAN内 IGMP Snooping的版本为 3，才能支持SSM Mapping功能。但如果配置了组播VLAN复制功能，只需在组播VLAN内配置SSM Mapping即可。

说明

虽然配置SSM-Mapping时，需要在VLAN下指定IGMP的版本号为3，但是在向路由器端口转发所收到的Version 2的 IGMP协议报文时并不会将其转换为Version 3版本。此时可以通过在交换机上配置 IGMP Snooping Proxy功能将其转换为Version 3的协议报文向上游发送。

配置 IGMP Snooping SSM Mapping的步骤如表 13-26所示。

表13-26 IGMP Snooping SSM Mapping的配置步骤

步骤	命令	说明
1	**system-view** 例如：<HUAWEI> **system-view**	进入系统视图
2	**vlan** *vlan-id* 例如：[HUAWEI]**vlan** 10	键入要配置 IGMP Snooping SSM Mapping 的 VLAN，进入 VLAN 视图
3	**igmp-snooping version 3** 例如：[HUAWEI-vlan10] **igmp-snooping version 3**	配置 VLAN 内 IGMP Snooping 的版本号为 3。缺省版本号为 2，但是 IGMPv2 版本不支持 SSM Mapping 功能
4	**igmp-snooping ssm-mapping enable** 例如：[HUAWEI-vlan10] **igmp-snooping ssm-mapping enable**	使能 VLAN 内的 SSM Mapping 功能。缺省情况下，VLAN 内 SSM Mapping 功能未使能，可用 **undo igmp-snooping ssm-mapping enable** 命令去使能 SSM Mapping 功能
5	**igmp-snooping ssm-mapping** *group-address* { *group-mask* | *mask-length* } *source-address* 例如：[HUAWEI-vlan10] **igmp-snooping ssm-mapping 238.1.1.0 24 10.1.1.1**	配置 VLAN 内组播组与组播源的映射。命令中的参数说明如下。 （1）*group-address*：指定要映射的组播组 IP 地址，缺省为 SSM 组策略范围内的组播组地址，也可以是前面 SSM 组策略中指定的组播组 IP 地址 （2）*group-mask*：二选一参数，组播组 IP 地址掩码 （3）*mask-length*：二选一参数，组播组 IP 地址掩码长度 （4）*source-address*：以上组播组 IP 地址要建立映射的组播源 IP 地址，是单播 IP 地址 缺省情况下，没有配置任何组播组与组播组源的映射，可用 **undo igmp-snooping ssm-mapping** *group-address* { *group-mask* | *mask-length* } *source-address* 命令取消对应 VLAN 中配置的指定组播组与组播源映射

【示例 2】配置设备上VLAN10中组播地址238.1.1.1～238.1.1.255与组播组源地址10.1.1.1之间的映射功能。

```
<HUAWEI>system-view
[HUAWEI] igmp-snooping enable
[HUAWEI] vlan 10
[HUAWEI-vlan10] igmp-snooping enable
[HUAWEI-vlan10] igmp-snooping version 3
[HUAWEI-vlan10] igmp-snooping ssm-mapping enable
[HUAWEI-vlan10] igmp-snooping ssm-mapping 238.1.1.0 24 10.1.1.1
```

13.4.10 IGMP Snooping管理

在日常维护工作中，可以在任意视图下选择执行以下 display 命令，了解 IGMP Snooping的运行状况（主要包括查看配置信息、成员端口和路由器端口、报文统计计数、转发表信息等）。

（1）使用display igmp-snooping [vlan [vlan-id]] 命令查看所有或者指定VLAN内IGMP Snooping配置信息，包括缺省配置信息。

（2）使用display igmp-snooping [vlan [vlan-id]] configuration命令查看所有或者指定VLAN内 IGMP Snooping非缺省配置信息。

（3）使用display igmp-snooping port-info [vlanvlan-id [group-addressgroup-address]] [verbose] 命令查看所有或者指定VLAN和组播组中的摘要或详细（选择 verbose可选项时）成员端口信息。

（4）使用display igmp-snooping router-port vlan vlan-id命令查看指定VLAN中的路由器端口信息。

（5）使用 display igmp-snooping queriervlan [vlan-id] 命令查看所有或者指定VLAN中的 IGMP Snooping查询器信息。

（6）使用 display igmp-snooping statisticsvlan [vlan-id] 命令查看所有或者指定VLAN中的 IGMP Snooping的统计信息。

（7）使用display l2-multicast forwarding-modevlan [vlan-id] 命令查看所有或者指定VLAN中的二层组播的转发模式。

（8）使用display l2-multicast forwarding-tablevlan vlan-id [[source-address source- address] group-address { group-address |router-group }] 命令查看指定 VLAN内二层组播转发表信息。

在接口视图下执行undo l2-multicast static-group [source-address source-ip-address] group-addressgroup-ip-addressvlan { all | { vlan-id1 [to vlan-id2] } &<1-10> }命令取消接口静态加入组播组的配置。也可以通过以下命令批量取消接口上加入的组播组地址。

（1）undo l2-multicast static-group [source-address source-ip-address] group-address group-ip-address1 to group-ip-address2vlan vlan-id

（2）undo l2-multicast static-group [source-address source-ip-address] group-address all vlan { all | { vlan-id1 [to vlan-id2] } &<1-10> }

在用户视图下使用 reset igmp-snooping group {all | vlan { all | vlan-id [[source- address source-address] group-address group-address] } }命令清除所有或者指定 VLAN 中的动态组表项。在用户视图下使用 reset igmp-snooping statistics {all | vlan { vlan- id | all } }命令清除所有或者指定VLAN中的 IGMP Snooping统计信息。IGMP Snooping的统计信息主要包括VLAN内接收到的Report、Leave、Query等协议报文的数量，通过该命令可以将这些统计计数置0，便于重新统计。

13.4.11 IGMP Snooping基本功能配置示例

本示例拓扑结构如图 13-10 所示， Router通过二层设备Switch连接用户网络，Router上运行IGMPv2版本。组播源Source向组播组225.1.1.1～225.1.1.5发送数据，网络中有 HostA、HostB、HostC 三个组播组成员，它们只对225.1.1.1～225.1.1.3的数据感兴趣。

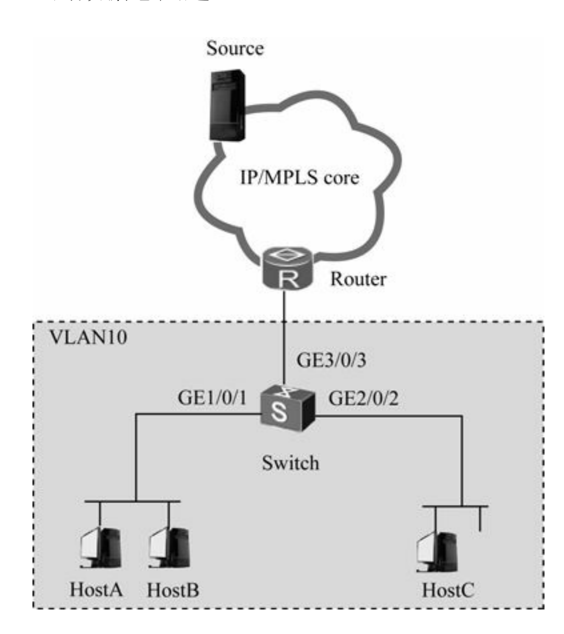


图13-10 IGMP Snooping基本功能配置示例拓扑结构

1．基本配置思路分析

因为示例中仅对VLAN 10中的用户进行组播组过滤，所以本示例只需要在使能 IGMP Snooping功能的基础上使用组策略就可以实现。具体配置任务如下。

（1）在Switch上创建VLAN并将接口加入对应的VLAN。

（2）使能全局和 VLAN 的 IGMP Snooping功能。

（3）配置组播组过滤策略，并在VLAN内应用此策略。

2．具体配置步骤

下面是本示例以上三项配置任务的具体配置步骤。

（1）创建VLAN，配置接口加入VLAN。

<HUAWEI>system-view

[HUAWEI] sysname Switch

[Switch] vlan 10

[Switch-vlan10] quit

[Switch] interfacegigabitethernet 1/0/1

[Switch-GigabitEthernet1/0/1] port hybrid untagged vlan 10

[Switch-GigabitEthernet1/0/1] port hybrid pvid vlan 10

[Switch-GigabitEthernet1/0/1] quit

[Switch] interfacegigabitethernet 2/0/2

[Switch-GigabitEthernet2/0/2] port hybrid untagged vlan 10

[Switch-GigabitEthernet2/0/2] port hybrid pvid vlan 10

[Switch-GigabitEthernet2/0/2] quit

[Switch] interfacegigabitethernet 3/0/3

[Switch-GigabitEthernet3/0/3] port hybrid pvid vlan 10

[Switch-GigabitEthernet3/0/3] port hybrid untagged vlan 10

[Switch-GigabitEthernet3/0/3] quit

（2）使能全局和VLAN的 IGMP Snooping功能。

[Switch] igmp-snooping enable

[Switch] vlan 10

[Switch-vlan10] igmp-snooping enable

[Switch-vlan10] quit

（3）通过基本 ACL配置组播组过滤策略，仅 VLAN 10中的用户接收组播地址为225.1.1.1～225.1.1.3的组播数据。

[Switch] acl 2000

[Switch-acl-basic-2000] rule permit source225.1.1.1 0

[Switch-acl-basic-2000] rule permit source 225.1.1.2 0

[Switch-acl-basic-2000] rule permit source225.1.1.3 0

[Switch-acl-basic-2000] quit

[Switch] vlan 10

[Switch-vlan10] igmp-snooping group-policy 2000

[Switch-vlan10] quit

配置好后，可使用display igmp-snooping port-info vlan 10命令查看Switch上的端口信息。具体如下。从中可看出组225.1.1.1～225.1.1.3已在Switch上动态生成成员端口GE1/0/1和GE2/0/2。

<Switch>display igmp-snooping port-info vlan 10

```
                    (Source, Group)   Port                       Flag
Flag: S:Static          D:Dynamic     M: Ssm-mapping
---------------------------------------------------------------------
VLAN 10, 3 Entry(s)
                    (*, 225.1.1.1)  GE1/0/1                     -D-
                                      GE2/0/2                     -D-
                                                 2 port(s)
                    (*, 225.1.1.2)  GE1/0/1                     -D-
                                      GE2/0/2                     -D-
                                                 2 port(s)
                    (*, 225.1.1.3)  GE1/0/1                       -D-
                                      GE2/0/2                     -D-
                                                 2 port(s)
---------------------------------------------------------------------
```

然后可以通过display l2-multicast forwarding-table vlan 10命令查看Switch上二层组播转发表。具体如下，从中可看出转发表中只有225.1.1.1～225.1.1.3的组播数据，满足示例中的要求。

```
<Switch>display l2-multicast forwarding-table vlan10
VLAN ID : 10, Forwarding Mode : IP
----------------------------------------------------------------------
                         (Source, Group)      Interface                  Out-Vlan
----------------------------------------------------------------------
                           Router-port        GigabitEthernet3/0/3       10
                         (*, 225.1.1.1)       GigabitEthernet1/0/1       10
                                                  GigabitEthernet2/0/2    10
                                                  GigabitEthernet3/0/3    10
                         (*, 225.1.1.2)       GigabitEthernet1/0/1          10
                                                  GigabitEthernet2/0/2    10
                                                  GigabitEthernet3/0/3    10
                         (*, 225.1.1.3)       GigabitEthernet1/0/1          10
                                                  GigabitEthernet2/0/2    10
                                                  GigabitEthernet3/0/3    10
--------------------------------------------------------------------
Total Group(s) : 3
```

13.4.12 通过静态端口实现二层组播的配置示例

本示例拓扑结构如图13-11所示，路由器Router通过二层设备Switch连接用户网络， Router 的用户侧三层 VLANIF 接口配置了225.1.1.1～225.1.1.5 的 IGMP 静态组，没有运行IGMP协议。网络中有HostA、HostB、HostC、HostD 4个组播组成员，其中HostA和 HostB 希望长期稳定接收 225.1.1.1～225.1.1.3的数据，HostC和HostD希望长期稳定接收225.1.1.4～225.1.1.5的数据。

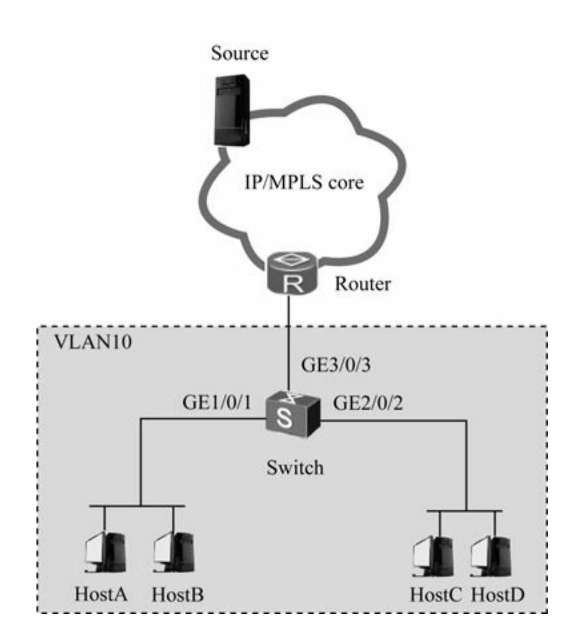


图13-11 静态端口配置示例拓扑结构

1．基本配置思路分析

本示例与上节介绍的示例要求差不多，但实现的方式不同。上节是通过组策略来实现用户仅接收来自指定组播组的数据，而本示例则要通过在组播组中添加静态端口（包括静态路由器端口和静态成员端口）的来实现仅接收来自指定组播组的数据。具体配置任务如下。

（1）在Switch上创建VLAN并将接口加入VLAN。

（2）使能全局和VLAN的 IGMP Snooping功能。

（3）配置静态路由器端口和配置静态成员端口。

2．具体配置步骤

（1）创建VLAN 10，并配置接口加入VLAN。参见上节配置。

（2）使能全局和VLAN IGMP Snooping功能。参见上节配置。

（3）把GE3/0/3接口配置为VLAN 10的静态路由器端口。

```
[Switch] interface gigabitethernet 3/0/3
[Switch-GigabitEthernet3/0/3] igmp-snooping static-router-port vlan 10
[Switch-GigabitEthernet3/0/3] quit
```

（4）把GE1/0/1和GE1/0/2接口分别加入到对应的组播组中，配置为它们的静态成员端口。

```
[Switch] interfacegigabitethernet 1/0/1
[Switch-GigabitEthernet1/0/1] l2-multicast static-group group-address 225.1.1.1 to 225.1.1.3 vlan 10
[Switch-GigabitEthernet1/0/1] quit
[Switch] interface gigabitethernet 2/0/2
[Switch-GigabitEthernet2/0/2] l2-multicast static-group group-address 225.1.1.4 to 225.1.1.5 vlan 10
[Switch-GigabitEthernet2/0/2] quit
```

配置好后，可以通过 display igmp-snooping router-port vlan 10 命令查看Switch上的路由器端口信息。具体如下，从中可以看出，GE3/0/3已成为静态路由器端口。

```
<Switch>display igmp-snooping router-port vlan 10
```

```
Port Name              UpTime        Expires        Flags
---------------------------------------------------------------------
VLAN 10, 1 router-port(s)
GE3/0/3               00:20:09             --             STATIC
```

同样可以通过display igmp-snooping port-info vlan 10命令查看Switch上的成员端口信息。具体如下，从中可看出组播组225.1.1.1～225.1.1.3在Switch上有静态成员端口GE1/0/1，组播组225.1.1.4～225.1.1.5在Switch上有静态成员端口GE2/0/2。

```
<Switch>display igmp-snooping port-info vlan 10
-----------------------------------------------------------------------
                      (Source, Group)    Port                  Flag
  Flag: S:Static       D:Dynamic       M: Ssm-mapping
-----------------------------------------------------------------------
VLAN 10, 5 Entry(s)
                      (*, 225.1.1.1) GE1/0/1                  S--
                                              1 port(s)
                      (*, 225.1.1.2) GE1/0/1                  S--
                                              1 port(s)
                      (*, 225.1.1.3) GE1/0/1                  S--
                                              1 port(s)
                      (*, 225.1.1.4) GE2/0/2                  S--
                                              1 port(s)
                      (*, 225.1.1.5) GE2/0/2                  S--
                                              1 port(s)
-----------------------------------------------------------------------
```

还可通过display l2-multicast forwarding-table vlan 10命令查看Switch上二层组播转发表。具体如下，从中可以看出，组 225.1.1.1～225.1.1.5 在 Switch 上已生成转发表。

```
<Switch>display l2-multicast forwarding-table vlan10
VLAN ID : 10, Forwarding Mode : IP
---------------------------------------------------------------------------
                           (Source, Group)    Interface                  Out-Vlan
---------------------------------------------------------------------------
                             Router-port       GigabitEthernet3/0/3          10
                           (*, 225.1.1.1)      GigabitEthernet1/0/1          10
                                                 GigabitEthernet3/0/3          10
                           (*, 225.1.1.2)      GigabitEthernet1/0/1          10
                                                 GigabitEthernet3/0/3          10
                           (*, 225.1.1.3)      GigabitEthernet1/0/1          10
                                                 GigabitEthernet3/0/3          10
                           (*, 225.1.1.4)      GigabitEthernet2/0/2          10
                                                 GigabitEthernet3/0/3          10
```

```
(*, 225.1.1.5)      GigabitEthernet2/0/2        10
                    GigabitEthernet3/0/3        10
-------------------------------------------------------------------------
Total Group(s) : 5
```

13.4.13 IGMP Snooping查询器的配置示例

本示例拓扑结构如图13-12所示，在一个没有三层设备纯二层网络环境中，组播源Source1和Source2分别向组播组224.1.1.1和225.1.1.1发送组播数据，HostA和HostC希望接收组播组224.1.1.1的数据，HostB和HostD希望接收组播组225.1.1.1的数据。所有组播组成员运行IGMPv2。

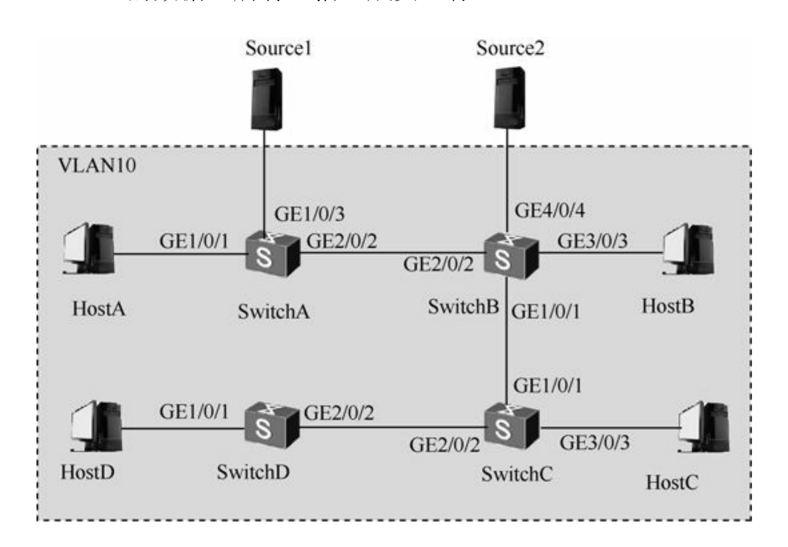


图13-12 IGMP Snooping查询器配置示例拓扑结构

本示例可通过在网络中各Switch上使能 IGMP Snooping功能，并配置某一台Switch为 IGMP Snooping查询器来实现。同时为防止设备在没有二层组播转发表项时将组播数据在VLAN内广播，在所有Switch上都使能丢弃未知组播报文功能。具体配置步骤如下。

（1）在所有Switch上创建VLAN并将接口加入VLAN。因为SwitchA、SwitchB、SwitchC、SwitchD的配置方法一样，现仅以SwitchA为例进行介绍。

```
<HUAWEI>system-view
[HUAWEI] sysname SwitchA
[SwitchA] vlan 10
[SwitchA-vlan10] quit
[SwitchA] interface gigabitethernet 1/0/1
[SwitchA-GigabitEthernet1/0/1] port hybrid pvid vlan 10
[SwitchA-GigabitEthernet1/0/1] port hybrid untagged vlan 10
[SwitchA-GigabitEthernet1/0/1] quit
[SwitchA] interface gigabitethernet 2/0/2
[SwitchA-GigabitEthernet2/0/2] port hybrid pvid vlan 10
[SwitchA-GigabitEthernet2/0/2] port hybrid untagged vlan 10
[SwitchA-GigabitEthernet2/0/2] quit
```

[SwitchA] interface gigabitethernet 3/0/3

[SwitchA-GigabitEthernet3/0/3] port hybrid pvid vlan 10

[SwitchA-GigabitEthernet3/0/3] port hybrid untagged vlan 10

[SwitchA-GigabitEthernet3/0/3] quit

（2）在所有Switch上使能全局和VLAN的 IGMP Snooping功能。同样因为SwitchA、SwitchB、SwitchC、SwitchD的配置方法一样，也仅以SwitchA为例进行介绍。

[SwitchA] igmp-snooping enable

[SwitchA] vlan 10

[SwitchA-vlan10] igmp-snooping enable

[SwitchA-vlan10] quit

（3）配置SwitchA为查询器。

[SwitchA] vlan 10

[SwitchA-vlan10] igmp-snooping querier enable

[SwitchA-vlan10] quit

（4）在所有Switch上使能丢弃未知组播报文功能。同样仅以SwitchA为例进行介绍。

[SwitchA] vlan 10

[SwitchA-vlan10] multicast drop-unknown

[SwitchA-vlan10] quit

当IGMP Snooping查询器开始工作之后，除查询器以外的所有设备都应能收到IGMP普遍组查询报文。可以通过display igmp-snooping statistics vlan 10命令查看 IGMP报文的统计信息，例如查看SwitchB上收到的IGMP报文统计信息。

```
<SwitchB>display igmp-snooping statistics vlan 10
IGMP Snooping Packets Counter
  Statistics for VLAN 10
    Recv V1 Report          0
    Recv V2 Report          32
    Recv V3 Report          0
    Recv V1 Query           0
    Recv V2 Query           30
    Recv V3 Query           0
    Recv Leave              0
    Recv Pim Hello          0
    Send Query(S=0)         0
    Send Query(S!=0)        0
    Suppress Report         0
    Suppress Leave          0
    Proxy Send General Query        0
    Proxy Send Group-Specific Query        0
    Proxy Send Group-Source-Specific Query 0
```

13.5 组播VLAN配置与管理

组播VLAN（Multicast VLAN）一般部署于设备的网络侧来实现组播流汇聚，然后将组播报文在用户VLAN内复制分发。华为S系列交换机支持基于用户VLAN和基于接口两种方式配置组播VLAN复制功能，可根据不同的应用场景来选择对应方式配置组播VLAN复制功能。有关组播VLAN工作原理参见本书第12章12.5.5节。

13.5.1 配置基于用户VLAN的组播VLAN一对多

通过配置组播VLAN一对多，可以实现组播数据在不同用户VLAN间复制分发，减少上游带宽浪费。目前S系列交换机在IPv4网络和IPv6网络都支持配置组播VLAN一对多。两种网络在配置时并无差异，都需要结合二层组播侦听功能（IPv4网络为IGMP Snooping，IPv6网络为MLD Snooping）来实现。在此仅以IPv4网络的配置流程进行介绍。具体配置顺序如下。

（1）配置用户VLAN IGMP Snooping功能

配置基于用户VLAN的组播VLAN一对多功能时需要在用户VLAN下使能二层组播侦听——IGMP Snooping功能。

（2）配置组播VLAN

组播 VLAN 是实现组播 VLAN 复制功能的基础，用来汇聚网络侧的组播流，然后将组播流在其对应的用户VLAN内复制分发。同时，在配置基于用户VLAN的组播VLAN功能时，组播VLAN也需要使能二层组播侦听功能。

（3）配置接口加入VLAN

组播VLAN和用户VLAN配置完成后，网络侧接口需要加入组播VLAN，用户侧接口需要加入用户VLAN。

注意

S2700EI/5700S-LI/5700LI系列交换机在配置组播VLAN时，用户侧接口必须以相同方式同时加入组播VLAN和用户VLAN。如果单个用户侧接口加入了两个或者两个以上的用户VLAN，只有第一个上送Report报文的用户VLAN才会实现组播VLAN复制功能。

以上基于用户 VLAN 的组播 VLAN 一对多的三项配置任务的具体配置步骤如表13-27所示。

表13-27 基于用户VLAN的组播VLAN一对多的配置步骤

配置任务	步骤	命令	说明
公共配置	1	**system-view** 例如：<HUAWEI> **system-view**	进入系统视图
配置用户 VLAN IGMP Snooping 功能	2	**igmp-snooping enable** 例如：[HUAWEI] **igmp-snooping enable**	使能全局 IGMP Snooping 功能
	3	**vlan** *vlan-id* 例如：[HUAWEI] **vlan** 10	创建要使能 IGMP Snooping 功能的用户 VLAN，并进入 VLAN 视图
	4	**igmp-snooping enable** 例如：[HUAWEI-vlan10] **igmp-snooping enable**	使能用户 VLAN 的 IGMP Snooping 功能
为所有需要接收组播数据的用户 VLAN 进行以上配置			
配置组播 VLAN	5	**quit** 例如：[HUAWEI-vlan10] **quit**	退出以上用户 VLAN 视图，返回系统视图
	6	**vlan** *vlan-id* 例如：[HUAWEI] **vlan** 5	创建要使能 IGMP Snooping 功能的组播 VLAN，并进入组播 VLAN 视图
	7	**igmp-snooping enable** 例如：[HUAWEI-vlan5] **igmp-snooping enable**	使能组播 VLAN 的 IGMP Snooping 功能
	8	**multicast-vlan enable** 例如：[HUAWEI-vlan5] **multicast-vlan enable**	使能组播 VLAN 功能，将当前 VLAN 配置为组播 VLAN 【注意】配置为组播 VLAN 的 VLAN，不能再被配置为用户 VLAN。被配置为用户 VLAN 的 VLAN 也不能再被配置为组播 VLAN；如果在 VLAN 内使用 **l2-multicast forwarding-mode** 命令将当前二层组播数据转发模式设置成了按 MAC 转发，则不支持将该 VLAN 配置为组播 VLAN。将组播 VLAN 恢复成普通 VLAN 时，应当先删除组播 VLAN 下的所有用户 VLAN 缺省情况下，当前 VLAN 为普通 VLAN，可用 **undo multicast-vlan enable** 命令将当前 VLAN 恢复成普通 VLAN
	9	**multicast-vlan user-vlan** { *vlan-id1* [**to** *vlan-id2*] } &<1-10> 例如：[HUAWEI-vlan5] **multicast-vlan user-vlan** 10 **to** 15	配置组播 VLAN 和用户 VLAN 的对应关系，将用户 VLAN 绑定到组播 VLAN。参数 *vlan-id1* [**to** *vlan-id2*] 用来指定要绑定组播 VLAN 的用户 VLAN 范围，取值范围均为 1～4 094 的整数 【说明】配置组播 VLAN 和用户 VLAN 的对应关系时，一个组播 VLAN 最多可以绑定 4 093 个用户 VLAN，而且所对应的用户 VLAN 必须已经创建，否则该命令即使配置成功也不生效，且一个用户 VLAN 只能绑定到一个组播 VLAN 缺省情况下，组播 VLAN 没有对应的用户 VLAN，可用 **undo multicast-vlan user-vlan** 命令取消组播 VLAN 和指定用户 VLAN 的对应关系
配置接口加入 VLAN	10	**quit** 例如：[HUAWEI-vlan5] **quit**	退出组播 VLAN 视图，返回系统视图
	11	将网络侧接口以 Trunk 方式或 Hybrid 方式加入组播 VLAN；将用户侧接口以 Trunk 方式或 Hybrid 方式加入用户 VLAN，具体参见本书第 6 章	

【示例 】配置组播VLAN2和用户VLAN3～VLAN10的对应关系。

<HUAWEI> system-view

[HUAWEI] vlan 2

[HUAWEI-vlan2] multicast-vlan enable

[HUAWEI-vlan2] multicast-vlan user-vlan 3 to 10

13.5.2 配置基于接口的组播**VLAN**功能

通过配置基于接口的组播 VLAN 功能，可以实现同一用户 VLAN 中不同用户之间的组播业务隔离，增强了对组播业务流量的控制。目前交换机仅支持在IPv4网络配置基于接口的组播VLAN功能。在配置时需要结合 IGMP Snooping功能来实现，但是与配置基于用户VLAN的组播VLAN功能不同的是，用户**VLAN**不需要使能 **IGMP Snooping**功能 ，只需使用vlan vlan-id命令创建用户VLAN。具体按如下顺序进行配置。

（1）配置组播VLAN

配置基于接口的组播 VLAN 功能时，只需要在组播 VLAN 下使能二层组播侦听功能，不需要使能组播VLAN功能。

（2）配置用户VLAN绑定组播VLAN

用户 VLAN 绑定组播 VLAN 主要在用户侧接口下进行配置，并且在同一接口下用户VLAN不能绑定到多个组播VLAN。

（3）配置接口加入VLAN

组播VLAN和用户VLAN配置完成后，网络侧接口需要加入组播VLAN，用户侧接口需要加入用户VLAN。

以上基于接口的组播VLAN的三项配置任务的具体配置步骤如表13-28所示。

表13-28 基于接口的组播VLAN的配置步骤

配置任务	步骤	命令	说明
公共配置	1	**system-view** 例如：<HUAWEI> **system-view**	进入系统视图
配置组播VLAN	2	**igmp-snooping enable**	使能全局 IGMP Snooping 功能
	3	**vlan** *vlan-id* 例如：[HUAWEI] **vlan** 5	创建要使能 IGMP Snooping 功能的组播 VLAN，并进入组播 VLAN 视图
	4	**igmp-snooping enable** 例如：[HUAWEI-vlan5] **igmp-snooping enable**	使能组播 VLAN 的 IGMP Snooping 功能
配置用户VLAN绑定组播VLAN	5	**quit** 例如：[HUAWEI-vlan5] **quit**	退出组播 VLAN 用户视图，返回系统视图
	6	**interface** *interface-type interface-number* 例如[HUAWEI] **interface gigabitethernet** 1/0/1	键入要绑定组播 VLAN 的用户侧**物理接口**，进入接口视图
	7	**l2-multicast-bind vlan** *vlanid1*[**to** *vlanid2*] **mvlan** *mvlanid* 例如：[HUAWEI-GigabitEthernet1/0/1] **l2-multicast-bind vlan** 100 **mvlan** 5	在以上物理接口下配置用户 VLAN 绑定组播 VLAN。命令中的参数说明如下。 （1）*vlan-id1* [**to** *vlan-id2*]：用来指定要绑定组播 VLAN 的用户 VLAN 范围，取值范围均为 1～4 094 的整数 （2）*mvlanid*：指定要绑定的组播 VLAN 缺省情况下，接口下没有配置用户 VLAN 绑定组播 VLAN，可用 **undo l2-multicast-bind vlan** 命令恢复缺省配置
配置接口加入VLAN	8	**quit** 例如：[HUAWEI-vlan5] **quit**	退出组播 VLAN 视图，返回系统视图
	9	将网络侧接口以 Trunk 方式或 Hybrid 方式加入组播 VLAN；将用户侧接口以 Trunk 方式或 Hybrid 方式加入用户 VLAN，具体参见本书第 6 章	

（续表）

【示例 】在接口GE1/0/1下配置用户VLAN100绑定组播VLAN20。

```
<HUAWEI> system-view
[HUAWEI] igmp-snooping enable
[HUAWEI] vlan 100
[HUAWEI-vlan100] igmp-snooping enable
[HUAWEI-vlan100] quit
[HUAWEI] vlan 20
[HUAWEI-vlan20] igmp-snooping enable
[HUAWEI-vlan20] quit
[HUAWEI] interfacegigabitethernet 1/0/1
[HUAWEI-GigabitEthernet1/0/1] l2-multicast-bind vlan 100mvlan 20
```

13.5.3 基于用户VLAN的组播VLAN配置示例

本示例拓扑结构如图13-13所示，RouterA和SwitchA之间用于传输组播数据的业 务 VLAN为 VLAN 10，而下游用户主机HostA、HostB 和 HostC 分别属于 VLAN 100、VLAN 200和VLAN 300，并且都需要接

收组播Source的组播数据。现要求通过配置基于用户 VLAN 的组播 VLAN 功能，满足对于不同用户主机有多份相同的组播需求。

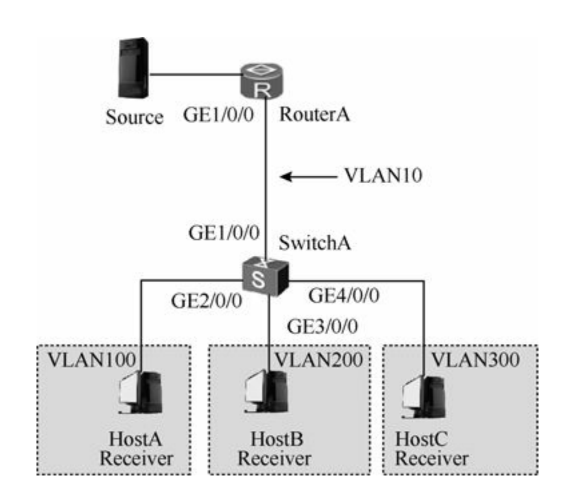


图13-13 基于用户VLAN的组播VLAN配置示例拓扑结构

1．基本配置思路分析

本示例可采用基于用户VLAN的组播VLAN功能来实现，基本配置思路如下。

（1）在 SwitchA 上使能全局的 IGMP Snooping功能。

（2）创建用户VLAN，并在用户VLAN下使能 IGMP Snooping。

（3）创建组播VLAN，并在组播VLAN下使能 IGMP Snooping。

（4）在组播VLAN下面绑定用户VLAN。

（5）将对应的接口分别以Hybrid方式加入对应的用户VLAN中。

2．具体配置步骤

（1）在系统视图下使能全局的 IGMP Snooping功能。

```
<SwitchA>system-view
[SwitchA] igmp-snooping enable
```

（2）创建用户VLAN，并在各用户VLAN下使能 IGMP Snooping功能。

```
[SwitchA] vlan 100
[SwitchA-vlan100] igmp-snooping enable
[SwitchA-vlan100] quit
[SwitchA] vlan 200
[SwitchA-vlan200] igmp-snooping enable
[SwitchA-vlan200] quit
[SwitchA] vlan 300
[SwitchA-vlan300] igmp-snooping enable
[SwitchA-vlan300] quit
```

（3）创建组播VLAN，并在组播VLAN下使能 IGMP Snooping功能。

```
[SwitchA] vlan 10
[SwitchA-vlan10] igmp-snooping enable
[SwitchA-vlan10] multicast-vlan enable
```

（4）在组播VLAN10下面绑定用户VLAN 100、VLAN 200和VLAN 300。

```
[SwitchA-vlan10] multicast-vlan user-vlan 100 200 300
[SwitchA-vlan10] quit
```

（5）把 GE1/0/0、GE2/0/0、GE3/0/0 和 GE4/0/0 接口以 Hybrid 方式加入对应的VLAN中。

```
[SwitchA] interfacegigabitethernet 1/0/0
[SwitchA-GigabitEthernet1/0/0] port hybrid pvid vlan 10
[SwitchA-GigabitEthernet1/0/0] port hybrid untagged vlan 10
[SwitchA-GigabitEthernet1/0/0] quit
[SwitchA] interfacegigabitethernet 2/0/0
[SwitchA-GigabitEthernet2/0/0] port hybrid pvid vlan 100
[SwitchA-GigabitEthernet2/0/0] port hybrid untagged vlan 100
[SwitchA-GigabitEthernet2/0/0] quit
[SwitchA] interfacegigabitethernet 3/0/0
[SwitchA-GigabitEthernet3/0/0] port hybrid pvid vlan 200
[SwitchA-GigabitEthernet3/0/0] port hybrid untagged vlan 200
[SwitchA-GigabitEthernet3/0/0] quit
[SwitchA] interface gigabitethernet 4/0/0
[SwitchA-GigabitEthernet4/0/0] port hybrid pvid vlan 300
[SwitchA-GigabitEthernet4/0/0] port hybrid untagged vlan 300
[SwitchA-GigabitEthernet4/0/0] quit
```

配置好后，可在SwitchA上使用display multicast-vlan vlan命令查看到组播VLAN和用户VLAN的信息。

```
[SwitchA] display multicast-vlan vlan
Total multicast vlan       1
multicast-vlan       user-vlan number       snooping-state
-----------------------------------------------------------------
10                         3                       IGMP Enable /MLD Disable
[SwitchA] display user-vlan vlan
   Total user vlan       3
user-vlan          snooping-state                      multicast-vlan          snooping-state
-----------------------------------------------------------------------------
100              IGMP Enable /MLD Disable       10                    IGMP Enable /MLD Disable
200              IGMP Enable /MLD Disable       10                    IGMP Enable /MLD Disable
300              IGMP Enable /MLD Disable       10                    IGMP Enable /MLD Disable
```

13.5.4 基于接口的组播VLAN配置示例

本示例拓扑结构如图13-14所示，SwitchA上的GE1/0/0接口连接路由器，GE2/0/0和GE3/0/0接口下的业务分别批发给 ISP1和 ISP2，ISP1和 ISP2分别通过组播VLAN 2和组播VLAN 3传输组播数据。GE2/0/0和GE3/0/0接口下用户VLAN重复，都为VLAN 10。为了防止不同ISP的组播报文会发送到不属于此ISP的用户，影响到ISP的利益，现要求通过基于接口的组播VLAN功能，指定属于本ISP的组播数据只转发到连接本ISP用户的接口。

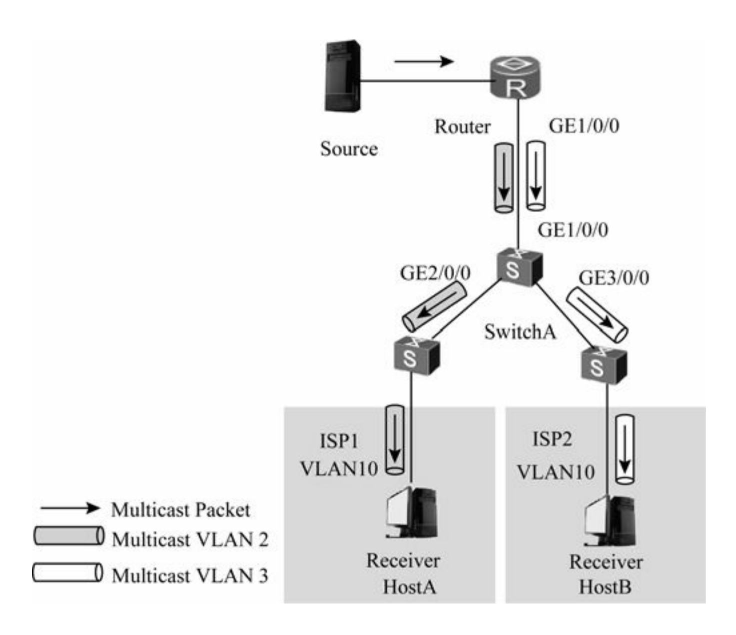


图13-14 基于接口的组播VLAN配置示例的拓扑结构

1．基本配置思路分析

本示例可采用基于接口的组播VLAN功能来实现，基本的配置思路如下所示。

（1）在系统视图下使能全局 IGMP Snooping功能。

（2）创建用户VLAN 10。

（3）创建组播VLAN 2和组播VLAN 3，并在组播VLAN下使能 IGMP Snooping。

（4）在GE2/0/0接口和GE3/0/0接口下对组播VLAN和用户VLAN分别进行绑定。

（5）将对应的接口分别以Hybrid方式加入VLAN。

2．具体配置步骤

（1）创建用户VLAN 10。

```
<SwitchA>system-view
[SwitchA] vlan 10
```

（2）配置组播VLAN 2和组播VLAN 3，并在组播VLAN下使能 IGMP Snooping功能。

```
[SwitchA] igmp-snooping enable
[SwitchA] vlan 2
[SwitchA-vlan2] igmp-snooping enable
[SwitchA-vlan2] quit
[SwitchA] vlan 3
[SwitchA-vlan3] igmp-snooping enable
[SwitchA-vlan3] quit
```

（3）在GE2/0/0和GE3/0/0接口下分别对组播VLAN和用户VLAN进行绑定。

```
[SwitchA] interface gigabitethernet 2/0/0
[SwitchA-GigabitEthernet2/0/0] l2-multicast-bind vlan 10 mvlan 2
[SwitchA-GigabitEthernet2/0/0] quit
[SwitchA] interface gigabitethernet 3/0/0
[SwitchA-GigabitEthernet3/0/0] l2-multicast-bind vlan 10 mvlan 3
```

[SwitchA-GigabitEthernet3/0/0] quit

（4）以Trunk方式把GE1/0/0接口加入组播VLAN 2和组播VLAN 3。

[SwitchA] interfacegigabitethernet 1/0/0

[SwitchA-GigabitEthernet1/0/0] port link-type trunk

[SwitchA-GigabitEthernet1/0/0] port trunk allow-pass vlan 2 3

[SwitchA-GigabitEthernet1/0/0] quit

（5）把GE2/0/0、GE3/0/0接口分别以Hybrid方式加入用户VLAN 10。

[SwitchA] interface gigabitethernet 2/0/0

[SwitchA-GigabitEthernet2/0/0] port hybrid pvid vlan 10

[SwitchA-GigabitEthernet2/0/0] port hybrid untagged vlan 10

[SwitchA-GigabitEthernet2/0/0] quit

[SwitchA] interface gigabitethernet 3/0/0

[SwitchA-GigabitEthernet3/0/0] port hybrid pvid vlan 10

[SwitchA-GigabitEthernet3/0/0] port hybrid untagged vlan 10

[SwitchA-GigabitEthernet3/0/0] quit

配置好后，可在 SwitchA上是使用 display l2-multicast-bind命令查看接口下用户VLAN与组播VLAN的绑定信息。具体如下。

```
[SwitchA] display l2-multicast-bind
-----------------------------------------------------------------
Port                    Startvlan          Endvlan        Mvlan
-----------------------------------------------------------------
GigabitEthernet2/0/0    10                      --          2
GigabitEthernet3/0/0    10                      --          3
-----------------------------------------------------------------
Total Table(s) : 2
```

第14章　镜像配置与管理

在日常的网络维护中经常遇到网络性能不正常现象（如网络很卡、丢包严重、频繁断网等），我们就会怀疑网络中有病毒在频繁发送广播报文，或者网络中某台主机遭到了不明的攻击，或者网络中有用户进行非法的网络应用（如网上看视频、下载大容量文件等）等。此时最有效的手段就是对网络中的特定用户、协议、端口、VLAN中的流量进行捕获，然后利用一些专门的工具软件进行分析。而这时首先必须要做的一件事就是在网络设备上配置好镜像功能，把要监控的流量复制一份到监控设备上，以便在监控设备捕获要监控的流量。

在华为S系列交换机中，“镜像”是这样描述的：在不影响报文正常处理流程的情况下，将镜像端口（源端口）的报文复制一份（并不是重定向原来的报文）到观察端口（目的端口），然后用户可以利用数据监控设备（如安装了Sniffer、科来、Wireshark等数据分析软件的设备）来分析复制到观察端口的报文，进行网络监控和故障排除。

在华为S系列交换机中，镜像又分为“端口镜像”、“流镜像”、“VLAN镜像”和“MAC地址镜像”四大类。而它们都有本地镜像和远程镜像之分，其中“端口镜像”和“流镜像”的远程镜像中又有二层远程镜像和三层远程镜像之分，而“VLAN镜像”和“MAC地址镜像”中的远程镜像只有二层远程镜像。而且不同的S系列交换机所支持的镜像类型不完全一样，本章将具体介绍各S系列交换机对这几种镜像特性的支持，以及配置与管理方法。

14.1 镜像基础

“镜像”是指将镜像端口（源端口）的报文复制一份到观察端口（目的端口），然后利用监控设备来观察、分析复制到观察端口上的报文，以实现网络监控和故障排除。它涉及两个重要的概念——镜像端口和观察端口。

“镜像端口”是指被监控的端口，也称镜像源端口，从镜像端口流经的所有指定方向或匹配流分类规则的报文将被复制到观察端口。而“观察端口”是指连接监控设备的端口，也称镜像目的端口，用于输出从镜像端口复制过来的报文，从而可以使用户监控到需要被监控的报文。

14.1.1 基本镜像原理

前面说了，镜像就是在不影响报文正常处理流程的情况下，把镜像端口上的报文复制一份到观察端口的过程，所以它不是报文的重定向，数据原来的传输路径仍然不会改变。根据镜像端口的数量不同，又分为“1∶1镜像”和“N∶1”镜像两大类。

1∶1镜像是指仅镜像一个端口上的报文到观察端口，即此时为一个镜像端口、一个观察端口。如图14-1所示，镜像端口B的报文被复制到观察端口C。

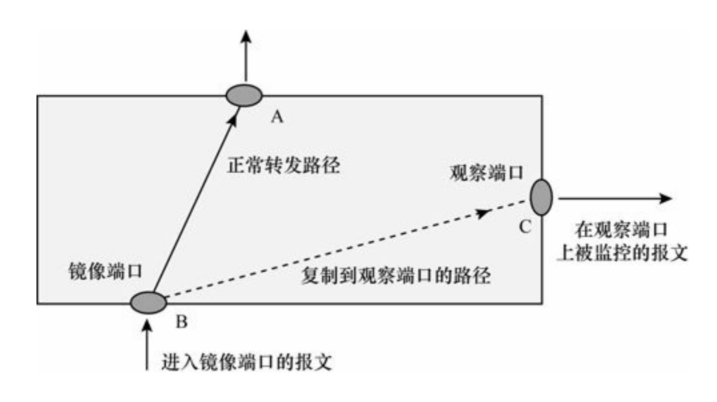


图14-1 1：1镜像示意图

N：1镜像是指镜像多个端口的报文到观察端口，即此时为多个镜像端口，一个观察端口，表示多个镜像端口上的报文可以镜像到同一个观察端口上。如图14-2所示，镜像端口B和镜像端口D的报文分别被复制了一份到观察端口C。

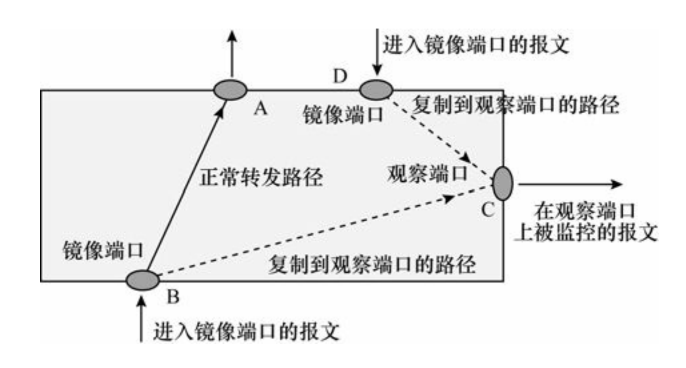


图14-2 N：1镜像示意图

14.1.2 镜像分类

在华为S系列交换机中，总体支持以下几种类型的镜像：端口镜像、流镜像、VLAN镜像和MAC地址镜像。但在华为S系列交换机的所有类型镜像中，**"**镜像端口**"**与**"**观察端口**"**只能在同一台交换机上配置 。且不同S系列交换机对这几类镜像的支持并不完全一样，下面分别予以介绍。

注意

除了端口镜像可以监控入方向或者出方向，或者同时监控入方向和出方向外，其 他类型镜像都仅可监控入方向的报文。当不再需要对报文进行监控时，建议取消镜像配置，以减少系统开销。

1．端口镜像

"端口镜像"也就是"基于端口的镜像"，是指复制一份从镜像端口流经的报文，然后传送到指定的观察端口的过程，如图14-3所示。

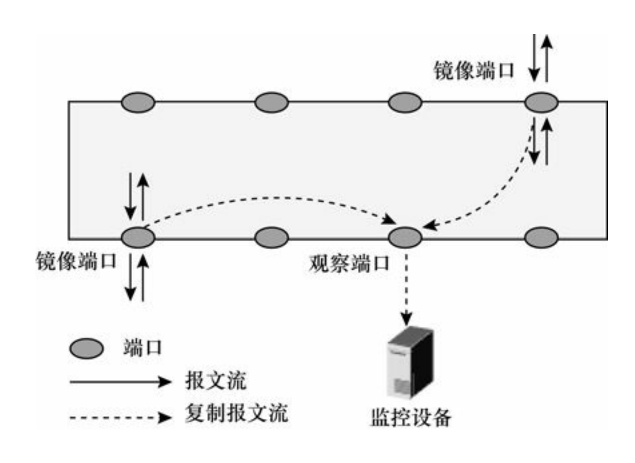


图14-3 端口镜像示意图

因为是基于端口的镜像，所以端口镜像可以监控的报文可以是任意方向，可以选择以下3种：

（1）入方向：仅对流入端口的报文进行镜像。

（2）出方向：仅对从端口上流出的报文进行镜像。

（3）双向：同时对流入端口和从端口上流出的报文进行镜像。

端口镜像分为本地端口镜像和远程端口镜像。“本地端口镜像”是指监控设备与观察端口直接相连，如图14-4所示。

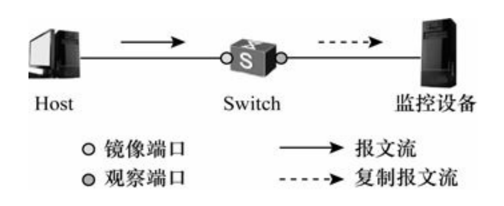


图14-4 本地镜像示意图

“远程端口镜像”是指监控设备与观察端口不是直接相连。在这里又分两种情况。

（1）二层远程端口镜像 RSPAN （ Remote Switched Port Analyzer，远程交换端口分析器）：监控设备与观察端口所连监控设备之间通过二层网络相连。被监控设备将流经镜像端口 的报文封装在镜像VLAN中，然后通过被监控设备的观察端口在远程镜像VLAN中广播，再将报文转发至监控设备，如图14-5所示。

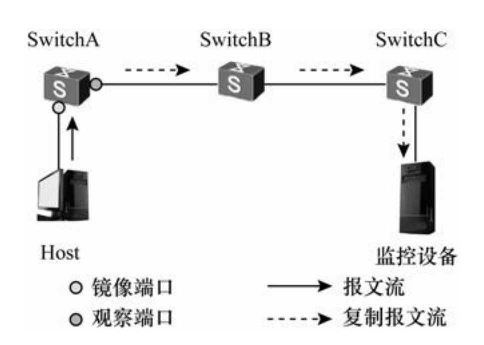


图14-5 二层远程端口镜像示意图

注意

在二层远程端口镜像中（其他镜像类型的 二层远程镜像也一样），**“观察端口”**和监控设备直接连接的

交换机端口都必须同时加入到镜像**VLAN**中，但**"**镜像端口**"**不能加入**VLAN**中 。 另外，从监控设备所连接的交换机到观察端口的所有二层设备相连的 Trunk 或者带标签的Hybrid端口必须允许镜像VLAN通过，以实现二层互通。

（2）三层远程端口镜像 ERSPAN （Encapsulated Remote SPAN，封装的远程交换端口分析器）：监控设备与观察端口所在设备之间通过三层网络相连。被监控设备将流经镜像端口的报文以GRE协议进行封装，然后GRE隧道通过三层IP网络传送到监控设备，如图14-6所示。隧道起始于被监控设备，终止于监控设备所连三层设备，GRE 报文源地址为被监控主机 IP 地址，目的地址为监控设备IP地址。当然首先要确保三层网络路由畅通。

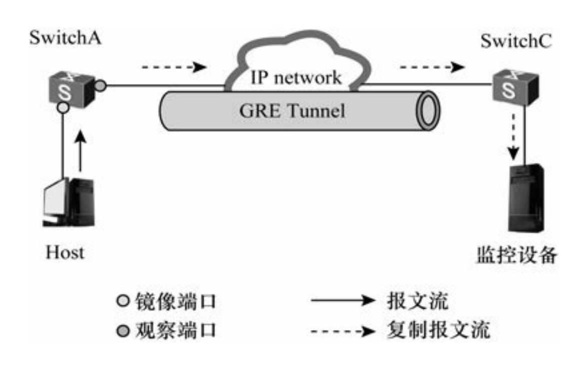


图14-6 三层远程端口镜像示意图

说明

流镜像也支持三层远程镜像，但与三层远程端口镜像一样，均仅 S7700/9300/9300E/9700系列交换机支持。

2．流镜像

"流镜像"也就是"基于流的镜像"，就是根据用户配置的策略，将镜像端口上指定的入方向（不支持出方向的流镜像） 报文复制到观察端口进行分析和监控。在流镜像中，要在镜像端口入方向应用包含流镜像行为的流策略。如果从镜像端口流经的报文匹配流分类规则，则将被复制到观察端口，如图14-7所示。

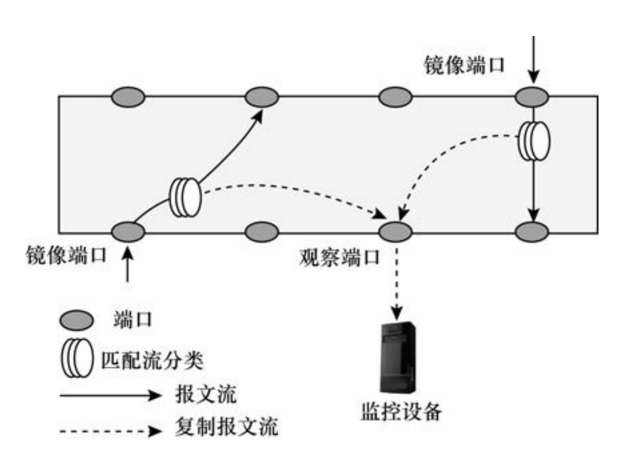


图14-7 流镜像示意图

流镜像也分本地流镜像和远程流镜像两大类，而远程流镜像也分二层远程流镜像和三层远程流镜像。本地流镜像、二层远程流镜像和三层远程流镜像的网络结构分别参见图14-4、图14-5和图14-6。但三层远程

流镜像**S7700/9300/9300E/9700**系列交换机支持 。

3．VLAN镜像

VLAN镜像也就是“基于VLAN的镜像”，是指将指定VLAN内所有活动接口的入方向（不支持出方向的**VLAN** 镜像） 上的报文镜像到观察端口。用户可以对某个VLAN或者某些VLAN内的报文进行监控。

VLAN镜像仅分本地VLAN镜像和二层远程VLAN镜像两类，无三层远程VLAN镜像。本地VLAN镜像、二层远程VLAN镜像的网络结构分别参见图14-4和图14-5。

4．MAC地址镜像

MAC 地址镜像即“基于 MAC 地址的镜像”，是指将匹配源或目的**MAC**地址的入方向（不支持出方向的**MAC**地址镜像） 报文镜像到观察端口。MAC地址镜像提供了一种更加精确的镜像 方式，用户可以对网络中特定设备的报文进行监控。

MAC地址镜像也仅分本地MAC地址镜像和二层远程MAC地址镜像两类，无三层远程MAC地址镜像。本地MAC地址镜像、二层远程MAC地址镜像的网络结构分别参见图14-4和图14-5。

14.1.3 镜像特性的产品支持

在上节介绍的4种镜像特性中，不同S系列交换机对它们的支持并不完全一样。下面分系列具体介绍。

1．S2700/3700系列交换机的镜像特性支持

（1）S2700SI和S2710SI系列不支持基于流、VLAN和MAC地址的远程镜像（但仍支持这些类型的本地镜像和基于端口的本地、远程镜像）。其他S2700/3700系列均同时支持基于端口、流、VLAN和MAC地址的本地和远程镜像。

（2）S2700和S3700系列均支持将多个端口的入方向和出方向报文镜像到一个观察端口，并且支持两个方向的报文镜像到同一观察端口。但不支持一条流同时镜像到多个观察接口 。

（3）S3700、S2710SI、S2700-52P-EI和S2700-52P-PWR-EI系列支持配置4个观察端口；其他S2700系列仅支持配置1个观察端口。

2．S5700/6700系列交换机的镜像特性支持

（1）S5700SI 不支持基于流、VLAN 和 MAC 地址的远程镜像（但仍支持这些类型的本地镜像和基于端口的本地、远程镜像）。其他系列均同时支持基于端口、流、VLAN和MAC地址的本地和远程镜像。

（2）S5700和S6700系列支持将多个端口的入方向和出方向报文镜像到一个观察端口，并且支持两个方向的报文镜像到同一观察端口。但不支持一条流同时镜像到多个观察接口 。

（3）S5700HI和S5710EI系列支持配置两个观察端口；S5700EI系列支持配置4个观察端口；S5700SI、S5700S-LI、S5700LI和S6700系列仅支持配置1个观察端口。

3．S7700/9300/9300E/9700系列交换机的镜像特性支持

（1）仅支持端口镜像和流镜像。

（2）设备支持跨板镜像，观察端口和镜像端口可以配置在同一台设备的不同接口板上。

（3）设备可以配置8个观察端口，支持将一个端口的报文镜像到多个观察端口，但镜像端口的一个方向的流量只能镜像到一个观察端口上 。

（4）设备支持将多个端口的入方向和出方向报文镜像到一个观察端口，并且支持两个方向的报文镜像到同一观察端口。

（5）设备每块单板最多可以同时运用3个观察端口，其中两个观察端口用于镜像端口入方向流量，一个观察端口用于镜像端口出方向流量。同一块单板的端口出方向流量只能镜像到同一个观察端口。但不支持对镜像报文进行再次镜像 。

14.2 端口镜像配置与管理

端口镜像是基于端口的镜像，可以对流经端口的入方向、出方向的报文进行单独镜像，或者两个方向的报文同时进行镜像。端口镜像又分本地端口镜像和远程端口镜像两种，而远程端口镜像中又分二层远程端口镜像和三层远程端口镜像两种。下面分别进行介绍。

配置好后，可通过 **display observe-port** 任意视图命令查看镜像的观察端口；通过**display port-mirroring** 任意视图命令查看镜像的配置信息。

14.2.1 配置本地端口镜像

通过配置本地端口镜像（网络结构参见图 14-4），可以将端口流经的报文复制到与本地观察端口连接的监控设备进行分析和监控。配置方法很简单，包括以下两项主要的配置任务。

1．配置本地观察端口

在本地镜像中，监控设备与观察端口是直接相连的，所以观察端口就是监控设备所连接的本地交换机上的端口。

2．配置镜像端口

镜像端口就是要被监控流量的端口，可以是以太网端口或 Eth-Trunk 端口，可以配置多个镜像端口到同一个观察端口的镜像 。如果配置Eth-trunk为镜像端口，则不能再单独配置其成员接口为镜像端口。相反，如果要配置Eth-trunk口下某成员接口为镜像端口，则不能再配置该成员对应的Eth-trunk口为镜像端口。

以上两项本地端口镜像配置任务的具体配置步骤如表14-1所示。当不再需要对报文进行监控时建议取消镜像配置，以减少系统开销。

表14-1 本地端口镜像配置步骤

配置任务	步骤	命令	说明
公共配置步骤	1	**system-view** 例如：<HUAWEI> **system-view**	进入系统视图
配置本地观察端口	2	**observe-port** *observe-port-index* **interface** *interface-type interface-number* 例如：[HUAWEI] **observe-port** 1 **interface** GigabitEthernet 2/0/0	配置本地观察端口。命令中的参数说明如下。 （1）*observe-port-index*：指定观察接口的索引，不同 S 系列交换机上的取值范围不同：S2700/5700SI/5700S-LI/5700LI/6700 系列只能为 1；S5700HI/5710EI 系列为 1～2 的整数；S23700/5700EI 系列为 1～4 的整数；S7700/9300/9300E/9700 系列为 1～8 的整数，且缺省值为 1 （2）*interface-type interface-number*：指定要作为观察端口的接口类型和编号。接口类型包括 Ethernet 接口（管理网口除外）、Eth-Trunk 接口、GigabitEthernet 接口和 XGE 接口 缺省情况下，系统没有配置本地观察端口，可用 **undo observe-port** *observe-port-index* 命令删除配置的指定号的本地观察端口。**如果要修改观察端口，则需先删除原来的配置后再重新配置**

（续表）

配置任务	步骤	命令	说明
配置镜像端口	3	**Interface** *interface-type interface-number* 例如：[HUAWEI] **interface** GigabitEthernet 2/0/10	键入要配置镜像端口的接口类型和编号。镜像端口可以是以太网端口或 Eth-Trunk 端口。建议观察端口和被观察端口同类型、同带宽
	4	**port-mirroring to observe-port** *observe-port-index* { **both** | **inbound** | **outbound** } 例如：[HUAWEI-GigabitEthernet2/0/10] **port-mirroring to observe-port** 1 **inbound**	配置基于端口的本地镜像。命令中的参数和选项说明如下。 （1）*observe-port-index*：指定要作为镜像端口的观察端口索引号，一定要与第 2 步配置的观察端口号一致 （2）**both**：多选一选项，指定同时监控以上端口的入和出两个方向报文 （3）**inbound**：多选一选项，指定仅监控以上端口的入方向报文 （4）**outbound**：多选一选项，指定仅监控以上端口的出方向报文 缺省情况下，接口不配置镜像功能，可用 **undo port-mirroring** [**to observe-port** *observe-port-index*] { **both** | **inbound** | **outbound** }命令取消对该接口的对应镜像功能

14.2.2 配置远程端口镜像

通过配置远程端口镜像，可以将端口流经的报文复制到远端监控设备进行分析和监控（网络结构参见图14-5和图14-6）。远程端口镜像又分二层远程端口镜像和三层远程端口镜像，但只有**S7700/9300/9300E/9700** 系列支持三层远程端口镜像。

远程端口镜像的主要配置任务如下。在配置远程端口镜像之前，需要确保观察端口所在设备与监控设备之间二层或三层网络互通。

1．配置远程观察端口

远程镜像中，监控设备与观察端口所在设备之间跨越二层或三层网络相连（网络结构分别参见图 14-5和图14-6）。设备将镜像报文封装VLAN或GRE IP报文，然后通过观察端口在远程镜像VLAN中广播，将报文转发至监控设备。

2．配置镜像端口

镜像端口可以是以太网端口或Eth-Trunk端口。同样，如果配置Eth-trunk口为镜像端口，则不能再单独配置其成员接口为镜像端口。相反，如果配置Eth-trunk口下某成员接口为镜像端口，则不能再配置Eth-trunk口为镜像端口。

远程端口镜像配置与上节介绍的本地镜像配置的唯一区别就在观察端口的配置上，镜像端口的配置方法是完全一样的。以上两项配置任务的具体配置步骤如表14-2所示。

表14-2 远程端口镜像的配置步骤

配置任务	步骤	命令	说明
公共配置步骤	1	**system-view** 例如：<HUAWEI> **system-view**	进入系统视图

（续表）

配置任务	步骤	命令	说明	
配置远程观察端口	2	**observe-port** *observe-port-index* **interface** *interface-type interface-number* **vlan** *vlan-id* 例如：[HUAWEI] **observe-port** 2 **interface** GigabitEthernet 2/0/0 **vlan** 10	（可选）配置**二层远程端口镜像**中的远程观察端口并指定远程镜像 VLAN。命令中的 *observe-port-index* 和 *interface-type interface-number* 参数与上节表 14-1 第 2 步中的说明一样，*vlan-id* 用来指定镜像报文封装的 VLAN ID（即镜像 VLAN），取值范围为 1～4 094。该 VLAN 需事先已创建好，并把观察端口加入到该 VLAN 中，**但镜像端口不允许加入远程镜像 VLAN** 缺省情况下，系统没有配置远程观察端口，可用 **undo observe-port** *observe-port-index* 命令删除配置的指定号的远程观察端口。**如果要修改观察端口，则需先删除原来的配置后再重新配置**	二选一
		observe-port [*observe-port-index*] **interface** *interface-type interface-number* **destination-ip** *dest-ip-address* **source-ip** *source-ip-address* [**dscp** *dscp-value* | **vlan** *vlan-id*] * 例如：[HUAWEI] **observe-port 2 interface** gigabitethernet 1/0/1 **destination-ip** 1.1.1.1 **source-ip** 2.2.2.2	（可选）配置**三层远程端口镜像**中的远程观察端口并指定远程镜像 VLAN。命令中的 *observe-port-index* 和 *interface-type interface-number* 参数也与上节表 14-1 第 2 步中的说明一样。其他参数说明如下。 （1）*dest-ip-address*：指定 GRE 报文的目的 IP 地址，是监控设备的 IP 地址 （2）*source-ip-address*：指定 GRE 报文的源 IP 地址，是被监控主机的 IP 地址 （3）*dscp-value*：可多选参数，指定 GRE 报文的优先级，取值范围为 0~63 的整数，缺省值为 0 （4）*vlan-id*：可多选参数，指定 GRE 隧道的二层报文封装的 VLAN ID。该 VLAN 需事先已创建好，并把观察端口加入该 VLAN 中，**但镜像端口不允许加入远程镜像 VLAN** 缺省情况下，系统没有配置远程观察端口，可用 **undo observe-port** *observe-port-index* 命令删除配置的指定号的远程观察端口。**如果要修改观察端口，则需先删除原来的配置后再重新配置**	
配置镜像端口	3	**Interface** *interface-type interface-number* 例如：[HUAWEI] **interface** GigabitEthernet 2/0/10	键入要配置镜像端口的接口类型和编号。镜像端口可以是以太网端口或 Eth-Trunk 端口。建议观察端口和被观察端口同类型、同带宽	
	4	**port-mirroring to observe-port** *observe-port-index* { **both** | **inbound** | **outbound** } 例如：[HUAWEI-GigabitEthernet2/0/10] **port-mirroring to observe-port** 2 **inbound**	配置基于端口的远程镜像，其他说明参见上节表 14-1 中的第 4 步	

【示例 1】配置GigabitEthernet1/0/2接口为二层远程镜像观察端口，接口所属VLAN为2。

<HUAWEI>system-view

[HUAWEI] observe-port 1 interface gigabitethernet 1/0/2vlan 2

【示例 2】配置GigabitEthernet1/0/1接口为三层远程镜像观察端口，目的IP为1.1.1.1，源IP为2.2.2.2。

<HUAWEI>system-view

[HUAWEI] observe-port 2 interface gigabitethernet 1/0/1destination-ip 1.1.1.1 source-ip 2.2.2.2

14.2.3 本地端口镜像配置示例

本示例拓扑结构如图14-8所示，HostA通过GigabitEthernet1/0/1接口接入SwitchA。监控设备Server直连在SwitchA的接口GigabitEthernet1/0/2上。用户希望通过监控设备 Server 对 HostA 发送的报文进行监控。

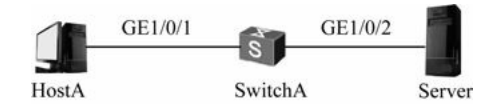


图14-8 本地端口镜像配置示例拓扑结构

本地端口镜像配置很简单，就是分别配置观察端口和镜像端口。本示例中，观察端口为SwitchA的GigabitEthernet1/0/1接口，镜像端口为SwitchA的GigabitEthernet1/0/2接口。下面是具体的配置步骤。

（1）在SwitchA上配置GigabitEthernet1/0/2接口为观察端口。

```
<HUAWEI>system-view
[HUAWEI] sysname SwitchA
[SwitchA] observe-port 1 interface gigabitethernet 1/0/2
```

（2）在SwitchA上配置GigabitEthernet1/0/1接口为镜像端口，以监控HostA发送的报文。

```
[SwitchA] interfacegigabitethernet 1/0/1
[SwitchA-GigabitEthernet1/0/1] port-mirroring to observe-port 1 inbound
[SwitchA-GigabitEthernet1/0/1] quit
```

配置好后，可以通过display observe-port任意视图命令查看观察端口的配置情况。具体如下，从中可以看到已配置的一个观察端口：GigabitEthernet1/0/2。

```
<SwitchA>display observe-port
----------------------------------------------------------------------------
  Index    : 1
  Interface: GigabitEthernet1/0/2
----------------------------------------------------------------------------
```

还可通过display port-mirroring任意视图命令查看镜像端口的配置情况。具体如下，从中可以看到已配置的镜像端口（Mirror-port）和观察端口（Observe-port），以及监控的报文方向为入方向（Inbound）。

14.2.4 二层远程端口镜像配置示例

本示例拓扑结构如图 14-9 所示， HostA 通过 GigabitEthernet1/0/2 接口接入SwitchA。监控设备Server接在SwitchC的GigabitEthernet1/0/1接口上。SwitchA与SwitchC通过二层网络互连。用户希望通过监控设备Server对HostA发送的报文进行远程监控。

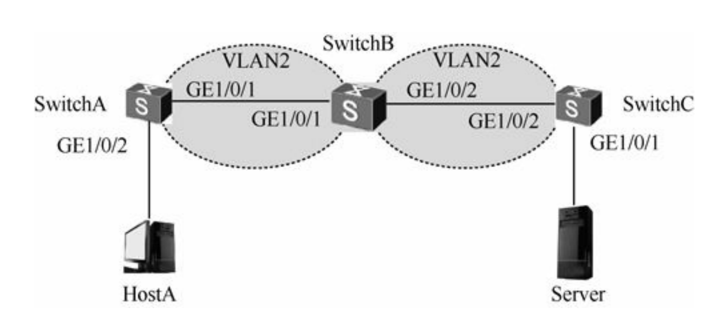


图14-9 二层远程端口镜像配置示例拓扑结构

二层远程端口镜像与本地端口镜像配置差不多，主要不同是需要事先配置好各二层交换机端口中的VLAN属性，以实现各二层设备间二层可达。另外在二层网络各交换机上创建镜像VLAN，把观察端口以及监控设备相连的交换机端口加入镜像VLAN中，同时要确保二层网络传输中各交换机相连接的Trunk端口允许镜像VLAN通过。具体配置步骤如下。

（1）配置各交换机接口类型，并按图中所示加入对应的VLAN中，使各设备间二层可达。

SwitchA上的配置：

<HUAWEI>system-view

[HUAWEI] sysname SwitchA

[SwitchA] vlan batch 2 to 3　#---批量创建VLAN 2和VLAN 3，其中VLAN 2是作为镜像VLAN的

[SwitchA] interface gigabitethernet 1/0/1

[SwitchA-GigabitEthernet1/0/1] port link-type trunk

[SwitchA-GigabitEthernet1/0/1] port trunk allow-pass vlan 2　#---在观察端口上允许镜像VLAN 2通过

[SwitchA-GigabitEthernet1/0/1] quit

[SwitchA] interface gigabitethernet 1/0/2

[SwitchA-GigabitEthernet1/0/2] port link-type access

[SwitchA-GigabitEthernet1/0/2] port default vlan 3　#---把镜像端口加入VLAN 3中

[SwitchA-GigabitEthernet1/0/2] quit

SwitchB上的配置：

<HUAWEI>system-view

[HUAWEI] sysname SwitchB

[SwitchB] vlan 2

[SwitchB-vlan2] quit

[SwitchB] interface gigabitethernet 1/0/1

[SwitchB-GigabitEthernet1/0/1] port link-type trunk

[SwitchB-GigabitEthernet1/0/1] port trunk allow-pass vlan 2　#---在二层设备间的链路上允许镜像VLAN 2通过

[SwitchB-GigabitEthernet1/0/1] quit

[SwitchB] interface gigabitethernet 1/0/2

[SwitchB-GigabitEthernet1/0/2] port link-type trunk

[SwitchB-GigabitEthernet1/0/2] port trunk allow-pass vlan 2　#---在二层设备间的链路上允许镜像VLAN 2通过

[SwitchB-GigabitEthernet1/0/2] quit

SwitchC上的配置：

<HUAWEI> system-view

[HUAWEI] sysname SwitchC

[SwitchC] vlan 2

[SwitchC-vlan2] quit

[SwitchC] interface gigabitethernet 1/0/1

[SwitchC-GigabitEthernet1/0/1] port link-type access　#---把监控设备相连端口加入镜像VLAN中

[SwitchC-GigabitEthernet1/0/1] port default vlan 2

[SwitchC-GigabitEthernet1/0/1] quit

[SwitchC] interface gigabitethernet 1/0/2

[SwitchC-GigabitEthernet1/0/2] port link-type trunk

[SwitchC-GigabitEthernet1/0/2] port trunk allow-pass vlan 2　#---在二层设备间的链路上允许镜像VLAN 2通过

[SwitchC-GigabitEthernet1/0/2] quit

（2）在SwitchA上配置GigabitEthernet1/0/1接口为远程观察端口，并且指定通过镜像VLAN 2广播。

[SwitchA] observe-port 1 interfacegigabitethernet 1/0/1vlan 2

（3）在SwitchA上配置GigabitEthernet1/0/2接口为镜像端口，并监控入方向的报文。

[SwitchA] interface gigabitethernet 1/0/2

[SwitchA-GigabitEthernet1/0/2] port-mirroring to observe-port 1 inbound

[SwitchA-GigabitEthernet1/0/2] quit

配置好后，可在 SwitchA上用 display observe-port任意视图命令查看观察端口的配置情况；用 display port-mirroring任意视图命令查看镜像端口的配置情况。结果如下。

```
<SwitchA>display observe-port
----------------------------------------------------------------------
  Index    : 1
  Interface: GigabitEthernet1/0/1
  Vlan    : 2
----------------------------------------------------------------------
<SwitchA>display port-mirroring
  Port-mirror:
----------------------------------------------------------------------
       Mirror-port               Direction     Observe-port
----------------------------------------------------------------------
1      GigabitEthernet1/0/2      Inbound       GigabitEthernet1/0/1
----------------------------------------------------------------------
```

14.2.5 三层远程端口镜像配置示例

如图14-10所示，HostA通过GigabitEthernet1/0/2接口接入SwitchA。监控设备Server连接在SwitchB的GigabitEthernet1/0/2接口上。HostA与Server之间跨越三层网络且路由可达。现用户希望通过监控设备 Server对 HostA 发送的报文进行远程监控。仅**S7700/9300/9300E/9700**系列交换机支持三层 远程端口镜像 。

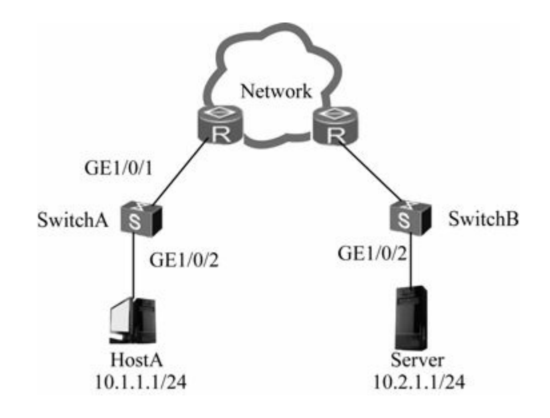


图14-10 三层远程端口镜像配置示例拓扑结构

本示例中假设已配置好了三层网络间的路由，三层远程端口镜像的配置就很简单了，只需要按照14.2.2

节中的表14-2所示的步骤配置好观察端口和镜像端口就行了。具体配置如下。

（1）在SwitchA上配置GigabitEthernet1/0/1接口为三层远程镜像观察端口，并指定封装的GRE报文报头中的目的IP地址为Server的IP地址，源IP地址为被监控主机HostA的IP地址。

```
<HUAWEI>system-view
[HUAWEI] sysname SwitchA
[SwitchA] observe-port 1 interface gigabitethernet 1/0/1destination-ip 10.2.1.1 source-ip 10.1.1.1
```

（2）在SwitchA上配置GigabitEthernet1/0/2接口为镜像端口，并指定仅监控该端口上的入方向报文。

```
[SwitchA] interfacegigabitethernet 1/0/2
[SwitchA-GigabitEthernet1/0/2] port-mirroring to observe-port1 inbound
[SwitchA-GigabitEthernet1/0/2] quit
```

配置好后，同样可使用 display observe-port任意视图命令查看观察端口的配置情况；使用display port-mirroring任意视图命令查看镜像端口的配置情况，验证配置结果。具体如下。

```
<SwitchA>display observe-port
----------------------------------------------------------------------
  Index    : 1
  Interface: GigabitEthernet1/0/1
  Vlan    : 0
  Dscp    : 0
  Src-Ip    : 10.1.1.1
  Dst-Ip    : 10.2.1.1
  Src_Mac    : 00-25-9e-37-ee-a5
  Dst_Mac    : ff-ff-ff-ff-ff-ff
----------------------------------------------------------------------
<SwitchA>display port-mirroring
  Port-mirror:
----------------------------------------------------------------------
    Mirror-port              Direction        Observe-port
----------------------------------------------------------------------
1   GigabitEthernet1/0/2   Inbound         GigabitEthernet1/0/1
----------------------------------------------------------------------
```

14.3 流镜像配置与管理

流镜像（即“基于流的镜像”）是通过QoS中的复杂流策略对特定的报文进行监控的方法（仅支持入方向的报文监控），也分本地流镜像和远程流镜像两类，而远程流镜像又分为二层远程流镜像和三层远程流镜像（同样仅**S7700/9300/9300E/9700**系列支持 ）。下面分别介绍这几种流镜像的配置与管理方法。

配置好后，可通过以下**display** 任意视图命令检查相关配置。

（1）使用**display observe-port** 命令查看镜像的观察端口。

（2）使用**display traffic behavioruser-defined** [behavior-name] 命令查看流镜像行为的配置信息。

（3）使用display **traffic classifieruser-defined** [classifier-name] 命令查看流分类的配置信息。

（4）执行**display traffic policy user-defined** [policy-name [**classifier** classifier-name]] 命令查看用户定义的流策略的配置信息。

（5）使用**display traffic-policy applied-record** [policy-name] 命令查看指定流镜像策略的应用记录信息。

14.3.1 配置本地流镜像

通过配置本地流镜像，可将镜像端口上流经的特定入方向 报文复制到本地的监控设备进行分析和监控。主要的配置任务如下。

（1）配置本地观察端口

本地镜像中，监控设备与观察端口直接相连。

（2）配置流分类

需根据实际应用，选择合适的流分类规则，定义对应的复杂流分类，具体配置参见本书第11章11.4.1节QoS流策略中的流分类。

（3）配置流镜像行为

定义将符合流分类规则的所有报文镜像到观察端口的流行为。具体配置参见本书第11章11.4.2节QoS流策略中的流行为。

（4）配置流镜像策略

创建一个流镜像策略，将以上定义的流分类和流行为关联起来。具体配置参见本书第11章11.4.3节QoS流策略中的流策略。

（5）应用流镜像策略

将关联了流行为与流分类的完整流策略应用到全局、接口或VLAN上。具体配置参见本书第11章11.4.4节QoS流策略中的流策略应用。

以上五项配置任务的具体配置步骤如表14-3所示。

表14-3 本地流镜像配置步骤

配置任务	步骤	命令	说明
公共配置步骤	1	**system-view** 例如：<HUAWEI> **system-view**	进入系统视图
配置本地流镜像观察端口	2	**observe-port** *observe-port-index* **interface** *interface-type interface-number* 例如：[HUAWEI] **observe-port** 1 **interface** GigabitEthernet 2/0/0	配置本地流镜像观察端口。其他说明参见表 14-1 中的第 2 步
配置流分类	3	根据本书第 11 章 11.4.1 节的介绍，选择适合的流分类方式对需要监控的报文定义一个流分类	
配置流行为	4	**traffic behavior** *behavior-name* 例如：[HUAWEI] **traffic behavior** b1	创建流镜像行为，并进入流镜像行为视图。参数 *behavior-name* 用来指定流行为名称，为 1～31 个字符，不支持空格，区分大小写 缺省情况下，系统未创建任何流行为，可用 **undo traffic behavior** *behavior-name* 命令删除指定的流镜像行为
	5	**mirroring to observe-port** *observe-port-index* 例如：[HUAWEI-behavior-b1] **mirroring to observe-port** 1	将满足规则的流镜像到指定的观察端口。参数 *observe-port-index* 用来指定观察端口索引号，一定要与第 2 步中指定的观察端口索引号一致 缺省情况下，系统未对任何报文定义流镜像动作，可用 **undo mirroring** 命令取消该流的镜像动作

（续表）

配置任务	步骤	命令	说明
公共配置步骤	1	**system-view** 例如：<HUAWEI> **system-view**	进入系统视图
配置本地流镜像观察端口	2	**observe-port** *observe-port-index* **interface** *interface-type interface-number* 例如：[HUAWEI] **observe-port** 1 **interface** GigabitEthernet 2/0/0	配置本地流镜像观察端口。其他说明参见表 14-1 中的第 2 步
配置流分类	3	根据本书第 11 章 11.4.1 节的介绍，选择适合的流分类方式对需要监控的报文定义一个流分类	
配置流行为	4	**traffic behavior** *behavior-name* 例如：[HUAWEI] **traffic behavior** b1	创建流镜像行为，并进入流镜像行为视图。参数 *behavior-name* 用来指定流行为名称，为 1～31 个字符，不支持空格，区分大小写 缺省情况下，系统未创建任何流行为，可用 **undo traffic behavior** *behavior-name* 命令删除指定的流镜像行为
	5	**mirroring to observe-port** *observe-port-index* 例如：[HUAWEI-behavior-b1] **mirroring to observe-port** 1	将满足规则的流镜像到指定的观察端口。参数 *observe-port-index* 用来指定观察端口索引号，一定要与第 2 步中指定的观察端口索引号一致 缺省情况下，系统未对任何报文定义流镜像动作，可用 **undo mirroring** 命令取消该流的镜像动作

14.3.2 配置远程流镜像

通过配置远程流镜像，可以将端口流经的特定入方向报文复制到远端的监控设备进行分析和监控。同样它可分二层远程流镜像和三层远程流镜像。在配置远程流镜像之前，需要确保观察端口所在设备与监控设备之间二层或三层网络互通。

远程流镜像的配置任务与上节介绍的本地流镜像配置任务一样，只是具体的配置步骤有些不同而已，具体如表14-4所示。

表14-4 远程流镜像配置步骤

配置任务	步骤	命令	说明	
公共配置步骤	1	**system-view** 例如：<HUAWEI> **system-view**	进入系统视图	
配置远程镜像观察端口	2	**observe-port** *observe-port-index* **interface** *interface-type interface-number* **vlan** *vlan-id* 例如：[HUAWEI] **observe-port** 2 **nterface** GigabitEthernet 2/0/0 **vlan** 10	（可选）配置**二层远程流镜像**观察端口。其他说明参见 14.2.2 节表 14-2 中的第 2 步对应命令参数说明	二选一
		observe-port [*observe-port-index*] **interface** *interface-type interface-number* **destination-ip** *dest-ip-address* **source-ip** *source-ip-address* [**dscp** *dscp-value* | **vlan** *vlan-id*] * 例如：[HUAWEI]**observe-port** 2 **interface** gigabitethernet 1/0/1 **destination-ip** 1.1.1.1 **source-ip** 2.2.2.2	（可选）配置**三层远程流镜像**观察端口。其他说明参见表 14-2 中的第 2 步对应命令参数说明	
配置流分类	3	根据本书第 11 章 11.4.1 节的介绍，选择适合的流分类方式对需要监控的报文定义一个流分类		
配置流行为	4	**traffic behavior** *behavior-name* 例如：[HUAWEI] **traffic behavior** b1	创建流镜像行为，并进入流镜像行为视图。其他说明参见上节表 14-3 中的第 4 步说明	
	5	**mirroring to observe-port** *observe-port-index* 例如：[HUAWEI-behavior-b1] **mirroring to observe-port** 1	将满足规则的流镜像到指定的观察端口。其他说明参见上节表 14-3 中的第 5 步说明	

（续表）

配置任务	步骤	命令	说明
配置流镜像策略	6	**quit** 例如：[HUAWEI-behavior-b1] **quit**	退出流行为视图，返回系统视图
	7	**traffic policy** *policy-name* 例如：[HUAWEI] **traffic policy** p1	创建流镜像策略，并进入流镜像策略视图。其他说明参见上节表 14-3 中的第 7 步说明
	8	**classifier** *classifier-name* **behavior** *behavior-name* 例如：[HUAWEI-trafficpolicy-p1] **classifier** c1 **behavior** b1	在流镜像策略中关联前面第 3 步和第 4 步配置的流分类和流镜像行为。其他说明参见上节表 14-3 中的第 8 步说明
应用流镜像策略	9	将关联了流行为与流分类的完整流策略应用到**镜像端口入方向**上，具体参见本书第 11 章 11.4.4 节	

14.3.3 本地流镜像配置示例

如图14-11所示，HostA通过GigabitEthernet1/0/1接口接入 SwitchA。监控设备 Server 直连在SwitchA的GigabitEthernet1/0/2接口上。现用户希望通过监控设备Server对HostA发出的802.1p优先级为6的报文进行监控。

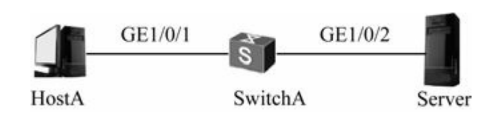


图14-11 本地流镜像配置示例拓扑结构

本示例最关键是要根据802.1p优先级进行流分类，然后定义远程流镜像行为、创建一个流策略，将定义的流分类和流行为进行关联，然后应用到镜像端口入方向上。具体配置如下。

（1）在SwitchA上配置GigabitEthernet1/0/2接口为观察端口。

```
<HUAWEI>system-view
[HUAWEI] sysname SwitchA
[SwitchA] observe-port 1 interface gigabitethernet 1/0/2
```

（2）在SwitchA上创建流分类c1，并配置流分类规则为匹配802.1p优先级为6的报文。

```
[SwitchA] traffic classifier c1
[SwitchA-classifier-c1] if-match 8021p 6
[SwitchA-classifier-c1] quit
```

（3）在SwitchA上创建流行为b1，并配置流镜像动作。

```
[SwitchA] traffic behaviorb1
[SwitchA-behavior-b1] mirroring to observe-port 1
[SwitchA-behavior-b1] quit
```

（4）在SwitchA上创建流策略p1，将以上定义的流分类c1和对应的流行为b1进行关联，然后将创建的流策略应用镜像端口——GigabitEthernet1/0/1 接口的入方向上，以实现对HostA发出的802.1p优先级为6的报文进行监控。

```
[SwitchA] traffic policy p1
[SwitchA-trafficpolicy-p1] classifier c1behavior b1
[SwitchA-trafficpolicy-p1] quit
[SwitchA] interface gigabitethernet 1/0/1
[SwitchA-GigabitEthernet1/0/1] traffic-policy p1 inbound
[SwitchA-GigabitEthernet1/0/1] quit
```

[SwitchA] quit

配置好后，可以通过display traffic classifier user-defined任意视图命令查看流分类的配置信息；通过display traffic policy user-defined任意视图命令查看流策略的配置信息，以验证配置结果。具体如下。

```
<SwitchA>display traffic classifier user-defined c1
  User Defined Classifier Information:
    Classifier: c1
      Precedence: 10
      Operator: OR   #---显示流分类工作模式为OR（或）模式
      Rule(s) : if-match 8021p 6   #---显示分类规则为匹配802.1p优先级为6的报文
<SwitchA>display traffic policy user-defined p1
  User Defined Traffic Policy Information:
  Policy: p1
    Classifier: c1
      Operator: OR   #---显示流策略工作模式为OR（或）模式
        Behavior: b1
          Mirroring to observe-port 1   #----显示流行为为镜像流量到索引号为1的观察端口
```

14.4 VLAN镜像配置与管理

通过 VLAN 镜像可将指定 VLAN 内所有活动接口的入方向 报文镜像到观察端口。用户可以对某个VLAN或者某些VLAN内的报文进行监控。同样它也分本地VLAN镜像和远程 VLAN 镜像，但是因为 VLAN镜像仅是二层网络中的镜像动作，所以只有二层网络远程VLAN镜像，没有三层远程VLAN镜像。也正如此，仅**S2700/3700/5700/6700**系列交换机支持**VLAN**镜像，**S7700/9300/9300E/9700**系列路由交换机不支持。

配置好后，可通过 display observe-port任意视图命令查看镜像的观察端口；通过display port-mirroring任意视图命令查看镜像的配置信息。

14.4.1 配置本地VLAN镜像

通过配置本地 VLAN 镜像，可以将某些 VLAN 中所有活动接口的入方向报文复制到本地的监控设备进行分析和监控。本地VLAN镜像的配置很简单，仅以下两项配置任务（当不再需要对报文进行监控时建议取消镜像配置，以减少系统开销）。

（1）配置本地观察端口

本地VLAN镜像中，监控设备与观察端口直接相连。

（2）配置VLAN镜像

指定要监控的VLAN，将VLAN中所有活动接口上入方向的二层报文镜像到观察端口。

以上两项本地VLAN镜像配置任务的具体配置如表14-5所示。

表14-5 本地VLAN镜像配置步骤

配置任务	步骤	命令	说明
公共配置步骤	1	**system-view** 例如：<HUAWEI> **system-view**	进入系统视图
配置本地观察端口	2	**observe-port** *observe-port-index* **interface** *interface-type interface-number* 例如：[HUAWEI] **observe-port** 1 **interface** GigabitEthernet 2/0/0	配置本地观察端口。其他说明参见 14.2.1 节表 14-1 中的第 2 步说明

（续表）

配置任务	步骤	命令	说明
配置 VLAN 镜像	3	**vlan** *vlan-id* 例如：[HUAWEI] **vlan** 10	键入要被监控的 VLAN，进入 VLAN 视图
	4	**mirroring to observe-port** *observe-port-index* **inbound** 例如：[HUAWEI-vlan10] **mirroring to observe-port** 2 **inbound**	配置基于 VLAN 的镜像动作。参数 *observe-port-index* 用来指定被镜像到的观察端口号，要与第 2 步所指定的索引号一致 缺省情况下，没有配置基于 VLAN 的镜像动作，可用 **undo mirroring** 命令取消基于该 VLAN 的镜像动作

14.4.2 配置远程VLAN镜像

通过配置远程 VLAN 镜像，可以对某些 VLAN 中所有活动接口的入方向 报文复制到远程监控设备进行分析和监控。在配置远程VLAN镜像之前，需要确保观察端口所在设备与监控设备之间二层网络互通。

远程VLAN镜像的配置也很简单，也就以下两项配置任务。

（1）配置远程观察端口

远程镜像中，监控设备与观察端口所在设备之间跨越二层网络相连。设备将镜像报文封装VLAN，然后通过观察端口在远程镜像VLAN中广播，将报文转发至监控设备。

（2）配置VLAN镜像

指定要监控的VLAN。

以上两项配置任务的具体配置步骤如表14-6所示。

表14-6 远程VLAN镜像配置步骤

配置任务	步骤	命令	说明
公共配置步骤	1	**system-view** 例如：<HUAWEI> **system-view**	进入系统视图
配置远程观察端口	2	**observe-port** *observe-port-index* **interface** *interface-type interface-number* **vlan** *vlan-id* 例如：[HUAWEI] **observe-port** 1 **interface** GigabitEthernet 2/0/0 **vlan** 10	配置远程观察端口并指定远程镜像 VLAN。其他说明参见 14.2.2 节表 14-2 中的第 2 步对应命令参数说明
配置 VLAN 镜像	3	**vlan** *vlan-id* 例如：[HUAWEI] **vlan** 10	键入要被监控的 VLAN，进入 VLAN 视图
	4	**mirroring to observe-port** *observe-port-index* **inbound** 例如：[HUAWEI-vlan10] **mirroring to observe-port** 2 **inbound**	配置基于 VLAN 的镜像动作。其他说明参见上节表 14-5 中的第 4 步

14.4.3 本地VLAN镜像配置示例

本示例拓扑结构如图 14-12 所示，HostA 和 HostB 通过 GigabitEthernet0/0/1 与GigabitEthernet0/0/2接口接入SwitchA，HostA和HostB同属于VLAN10。监控设备Server直连在 SwitchA 的 GigabitEthernet0/0/3 接口上。现用户希望通过监控设备 Server 对 VLAN10 的所有活动接口的入流量进行监控。

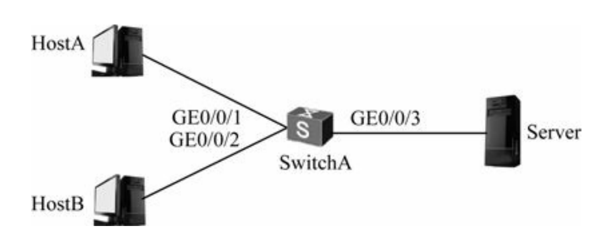


图14-12 本地VLAN镜像配置示例拓扑结构

根据 14.4.1 节介绍的配置方法可以很容易得出本示例具体配置步骤。但在此之前要先配置好各接口类型和所属VLAN。

（1）配置接口类型及所属VLAN。

```
<Switch>system-view
[Switch] sysname SwitchA
[SwitchA] vlan 10
[SwitchA-vlan10] quit
[SwitchA] interface gigabitethernet 0/0/1
[SwitchA-GigabitEthernet0/0/1] port link-type access
[SwitchA-GigabitEthernet0/0/1] port default vlan 10
[SwitchA-GigabitEthernet0/0/1] quit
[SwitchA] interface gigabitethernet 0/0/2
[SwitchA-GigabitEthernet0/0/2] port link-type access
[SwitchA-GigabitEthernet0/0/2] port default vlan 10
[SwitchA-GigabitEthernet0/0/2] quit
```

（2）配置GigabitEthernet0/0/3接口为观察端口。

```
[SwitchA] observe-port 1 interfacegigabitethernet 0/0/3
```

（3）配置监控VLAN10中所有活动接口入方向报文。

```
[SwitchA] vlan 10
[SwitchA-vlan10] mirroring to observe-port1 inbound
[SwitchA-vlan10] quit
```

配置好后可通过display port-mirroring任意视图命令查看VLAN镜像的配置情况，以验证配置结果。具体如下。

```
<SwitchA>display port-mirroring
  Vlan-mirror:
----------------------------------------------------------------------
  Mirror-vlan      Direction       Observe-port
----------------------------------------------------------------------
  10               Inbound         GigabitEthernet0/0/3
```

14.5 MAC地址镜像配置与管理

MAC地址镜像（即基于MAC地址的镜像）是将匹配源或目的MAC地址的入方向报文镜像到观察端口，提供了一种更加精确的镜像方式，用户可以对网络中特定设备的报文进行监控。它也有本地 MAC 地址镜像和远程 MAC 地址镜像两大类。同样因为它 是基于二层报文的报文镜像，所以在远程 MAC 地址镜像中也仅有二层远程 MAC 地址镜像这一种。但 **MAC** 地址镜像也仅 **S2700/3700/5700/6700** 系列交换机支持，**S7700/9300/9300E/9700**系列不支持 。

配置好后，可用display observe-port命令查看镜像的观察端口；可用display port-mirroring命令查看镜像的配置信息。

14.5.1 配置本地MAC地址镜像

通过配置本地 MAC 地址镜像，可以将指定源或目的 MAC 地址的报文复制到本地的监控设备进行分析和监控。其配置方法也很简单，仅需以下两项配置任务。

（1）配置本地观察端口

在本地镜像中，监控设备与观察端口直接相连。

（2）配置MAC地址镜像

指定用于匹配报文的源MAC地址或目的MAC地址，将符合条件的入方向的二层报文镜像到观察端口。

以上两项配置任务的具体配置步骤如表14-7所示。

表14-7 本地MAC地址镜像配置步骤

配置任务	步骤	命令	说明
公共 配置步骤	1	**system-view** 例如：<HUAWEI> **system-view**	进入系统视图
配置本地 观察端口	2	**observe-port** *observe-port-index* **interface** *interface-type interface-number* 例如：[HUAWEI] **observe-port** 1 **interface** GigabitEthernet 2/0/0	配置本地观察端口。其他说明参见 14.2.1 节表 14-1 中的第 2 步说明
配置 MAC 地址镜像	3	**vlan** *vlan-id* 例如：[HUAWEI] **vlan** 10	键入要被监控报文所在的 VLAN，进入 VLAN 视图
	4	**mac-mirroring** *mac-address* **to observe-port** *observe-port-index* **inbound** 例如：[HUAWEI-vlan10] **mac-mirroring** 1111-2222-3333 **to observe-port** 1 **inbound**	配置基于 MAC 地址的镜像。命令中的参数说明如下。 （1）*mac-address*：指定用于匹配报文的源 MAC 地址或者目的 MAC 地址（可匹配源 MAC 地址或目的 MAC 地址中的任意一个 MAC 地址） （2）*observe-port-index*：指定被镜像到的观察端口号，要与第 2 步所指定的索引号一致 缺省情况下，系统未配置 MAC 地址镜像，可用 **undo mac-mirroring** *mac-address* [**to observe-port** *observe-port-index*] **inbound** 命令删除 MAC 地址镜像

14.5.2 配置远程MAC地址镜像

通过配置远程MAC地址镜像，可以将指定源或目的MAC地址的报文复制到远程的监控设备进行分析和监控。远程MAC地址镜像只有二层远程MAC地址镜像，在配置二层远程MAC 地址镜像之前，需要确保观察端口所在设备与监控设备之间二层网络互通。

远程MAC地址镜像的配置任务与上节介绍的本地MAC地址镜像的配置任务一样，具体配置步骤如表14-8所示。

表14-8 本地MAC地址镜像配置步骤

配置任务	步骤	命令	说明
公共配置步骤	1	**system-view** 例如：<HUAWEI> **system-view**	进入系统视图
配置远程观察端口	2	**observe-port** *observe-port-index* **interface** *interface-type interface-number* **vlan** *vlan-id* 例如：[HUAWEI] **observe-port** 1 **interface** GigabitEthernet 2/0/0	配置远程观察端口。其他说明参见 14.2.2 节表 14-2 中的第 2 步对应命令参数说明
配置 MAC 地址镜像	3	**vlan** *vlan-id* 例如：[HUAWEI] **vlan** 10	键入要被监控的 VLAN，进入 VLAN 视图
	4	**mac-mirroring** *mac-address* **to observe-port** *observe-port-index* **inbound** 例如：[HUAWEI-vlan10] **mac-mirroring** 1111-2222-3333 **to observe-port** 1 **inbound**	配置基于 MAC 地址的镜像动作。其他说明参见 14.5.1 节表 14-7 中的第 4 步对应命令参数说明

14.5.3 本地MAC地址镜像配置示例

如图14-13所示，HostA和HostB通过GigabitEthernet0/0/1与GigabitEthernet0/0/2接口接入SwitchA，HostA和HostB同属于VLAN10。监控设备Server直连在SwitchA的 GigabitEthernet0/0/3 接口上。现用户希望对VLAN10 中源或目的 MAC 地址为0001-0001-0001的入方向流量进行监控。

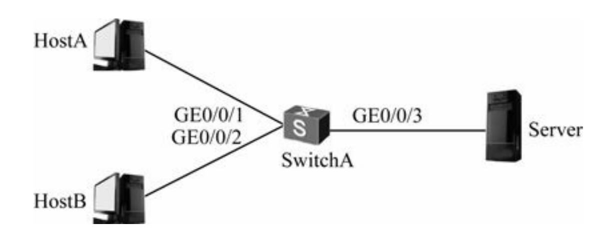


图14-13 本地MAC地址镜像配置示例拓扑结构

本地 MAC 地址镜像的配置很简单，本示例中只需配置 SwitchA 的GigabitEthernet0/0/3接口为观察端口，使直连的监控设备 Server 能够接收到镜像报文。然后在 VLAN10 视图下配置基于MAC地址的镜像动作即可。具体配置步骤如下。

（1）配置各接口类型和所属VLAN。

```
<Switch>system-view
[Switch] sysname SwitchA
[SwitchA] vlan 10
[SwitchA-vlan10] quit
[SwitchA] interfacegigabitethernet 0/0/1
[SwitchA-GigabitEthernet0/0/1] port link-type access
[SwitchA-GigabitEthernet0/0/1] port default vlan 10
[SwitchA-GigabitEthernet0/0/1] quit
[SwitchA] interface gigabitethernet 0/0/2
[SwitchA-GigabitEthernet0/0/2] port link-type access
```

[SwitchA-GigabitEthernet0/0/2] port default vlan 10

[SwitchA-GigabitEthernet0/0/2] quit

（2）配置GigabitEthernet0/0/3接口为观察端口。

[SwitchA] observe-port 1 interface gigabitethernet 0/0/3

（3）配置基于MAC地址的镜像动作。

[SwitchA] vlan 10

[SwitchA-vlan10] mac-mirroring 0001-0001-0001 to observe-port 1 inbound

[SwitchA-vlan10] quit

配置好后，可以通过display port-mirroring任意视图命令查看MAC地址镜像的配置情况，以验证配置结果。具体如下。

```
[SwitchA] display port-mirroring
  Mac-mirror:
----------------------------------------------------------------------
  Mirror-mac          Vlan         Direction       Observe-port
----------------------------------------------------------------------
  0001-0001-0001    10             Inbound         GigabitEthernet0/0/3
----------------------------------------------------------------------
```

第15章　基于MAC地址的安全配置与管理

MAC地址是网络设备中不变的物理地址，所以基于MAC地址的接入控制就成了最直接，甚至可能是大多数情况下最有效的控制手段。我们知道，在二层交换网络中，是通过依靠保存在交换机中的MAC地址表来进行寻址的，这时如果控制交换机中存储的MAC地址表就可以控制一些非法设备的接入，让其他设备不能与它进行通信。

在华为S系列交换机中，为了实现这个目的提供了多种基于MAC地址的安全手段，如限制接口学习MAC地址表项的数量，限制交换机中MAC地址表项的总数，关闭端口的MAC地址学习功能，使能端口安全功能（包括安全动态MAC地址和Sticky MAC地址）、MAC地址防漂移、MAC地址漂移检测、MAC-spoofing-defend、丢弃源MAC地址为全0或全1的报文和端口桥功能（丢弃源MAC地址和目的MAC地址均在设备的同一端口上学习到的报文）等。

本章将全面介绍以上几种华为S系列交换机中基于MAC地址的安全特性原理和具体的配置方法。但这些安全特性更适用于小型网络，或者对一些特定的非法用户进行控制，在大型网络中更多的是采用本书后面将要介绍的AAA控制方案和802.1x认证、MAC地址认证和Portal认证方式。

15.1 MAC地址表概述

MAC地址表记录了与交换机相连的设备的MAC地址、接口号以及所属的VLAN ID的对应关系，也就是通常所说的CAM（Content Addressable Memory，内容可寻址内存）表。在转发数据时，设备根据报文中的目的 MAC 地址查询 MAC 地址表，快速定位出接口，从而减少广播。

说明

MAC地址表与ARP表是不同的，ARP表中描述的是IP地址与MAC地址的对应 关系，用来通过IP地址解析MAC地址的，在局域网内部最终又是通过MAC地址表查找对应的出接口，进行数据帧转发的。

15.1.1 MAC地址表项

MAC地址表中的一个个映射项就是MAC地址表项，但它们有几种形成方式，对应就形成了几种MAC地址表项类型和老化方式。下面分别予以介绍。

1．MAC地址表的分类

MAC地址表项分为动态表项、静态表项和黑洞表项。

动态表项由接口通过对报文中的源MAC地址学习方式动态获得的，这类MAC地址表项有老化时间。静态MAC地址表项则是由用户手工配置的，这类MAC地址表项不会被老化。

黑洞MAC 地址表项是一种特殊的静态MAC 地址表项，用于丢弃含有特定源MAC地址或目的MAC地址的数据帧。为防止无用MAC地址表项占用MAC地址表，同时为了防止黑客通过MAC地址攻击用户设备或网络，可将非信任用户的MAC地址配置为黑洞MAC地址，当设备收到目的MAC或源MAC地址为黑洞MAC地址的报文时，直接丢弃。黑洞MAC地址表项也是由用户手工配置的，这类MAC地址表项也不会被老化。

在系统复位后，动态 MAC 地址表项会丢失，而保存的静态 MAC 地址表项和黑洞MAC地址表项都不会丢失。

2．MAC地址表项的生成方式

通过对前面MAC地址表项的类型介绍可以知道，MAC地址表项有两种生成方式：动态学习方式下的自动生成方式和手工配置方式。

（1）自动生成的MAC地址表项

一般情况下，MAC 地址表项是由设备通过对报文中源MAC 地址的学习而自动建立的。例如，当与SwitchA连接的SwitchB向SwitchA发送数据时，SwitchA从数据帧中解析出源MAC地址（即SwitchB的MAC地址），连同接口号添加到MAC地址表中。以后SwitchA接收到发送给SwitchB的数据，通过查询MAC表就可以得到正确的发送接口。

为适应网络的变化，MAC地址表项需要不断更新。MAC地址表项中自动生成的表项并非永久有效，每一条表项都有一个生存周期（也就是通常所说的“老化时间”），到达生存周期仍得不到刷新的表项将被删除。如果在到达生存周期前记录被刷新，则该表项的老化时间重新计算。

（2）手工配置的MAC地址表项

设备在通过报文中源 MAC 地址学习而自动建立 MAC 地址表项时无法区分合法用 户和黑客用户的报文，带来了安全隐患。如果黑客用户将攻击报文的源 MAC 地址伪装成合法用户的 MAC 地址，并从设备的其他接口进入，设备就会学习到错误的 MAC 地址表项，于是就会将本应转发给合法用户的报文转发给黑客用户。

为了提高接口安全性，网络管理员可手工在 MAC 地址表中加入静态 MAC 地址表项，相当于将用户MAC 地址与所连接的设备接口绑定，从而防止假冒身份的非法用户骗取数据。通过手工配置黑洞 MAC 地址表项，可以限制指定用户的流量不能从设备通过，防止非法用户的攻击。

手工配置的MAC表项优先级高于自动生成的表项。

3．基于MAC地址表的报文转发

设备在转发报文时，根据MAC地址表项信息，会采取以下两种转发方式。

（1）单播转发：当 MAC 地址表中包含与报文目的 MAC 地址对应的表项时，设备直接将报文从该表项中的转发出接口发送。

（2）广播转发：当设备收到的报文为广播报文（目的 MAC 地址为广播 MAC 地址ffff-ffff-ffff）、组播报文（目的MAC地址为组播MAC地址）或MAC地址表中没有包含对应报文目的 MAC 地址的表项时，设备将采取广播方式将报文向除接收接口外同一VLAN内的所有接口转发。

15.1.2 MAC地址表特性及产品支持

华为S系列交换机的MAC地址表特性主要包括MAC地址表基本功能和MAC地址表扩展功能两部分，来提高设备的安全性，控制MAC表的规模。MAC地址表基本功能如表15-1所示；MAC地址表扩展功能如表15-2所示。除非特殊说明，所有S系列交换机均支持这些功能。当然，在实际应用中，并不要求每项功能都配置，这些都是可根据实际需求选择配置的特性。

表15-1 MAC地址表基本功能

功能	说明
静态 MAC 地址表项	将一些固定的上行设备或信任用户的 MAC 地址配置为静态 MAC 表项，可以保证其安全通信
黑洞 MAC 地址表项	可以防止黑客通过 MAC 地址攻击网络
动态 MAC 地址表项老化时间	合理配置动态 MAC 表项的老化时间，可以防止 MAC 地址表项爆炸式增长
禁止 MAC 地址学习功能	适用于网络环境固定的场景或已经明确了转发路径的场景。可以限制非信任用户接入，防止 MAC 地址攻击，提高网络安全性。**S2700SI/27010SI 系列不支持**
限制 MAC 地址学习数量	在安全性比较差的网络中，可用于防止变换 MAC 地址攻击

表15-2 MAC地址表扩展功能

功能	说明
端口安全	在安全性要求较高的网络中，在端口上配置端口安全功能，可阻止其他非信任的主机通过本端口与设备通信，增强设备安全性。**S2700SI/27010SI/2700EI 系列不支持**
MAC-spoofing-defend 功能	配置该功能后，一个端口学习到的 MAC 地址在本设备其他端口上将不再学习，可以防止某些非法用户仿冒 MAC 地址来发送报文。**S5700HI/5710EI/700/9300/9300E/9700 系列不支持**

（续表）

功能	说明
MAC 地址防漂移	对于固定的上行设备或服务器，通过提高端口的优先级，可防止伪造 MAC 地址攻击。**S2700/3700/6700 系列不支持，在 S5700 系列中，也仅 S5700HI 和 S5710EI 支持**
MAC 地址漂移检测功能	配置 MAC 地址漂移检测功能可减少网络环路对本设备的影响。**S2700SI/27010SI/2700EI 系列不支持**
丢弃全零 MAC 地址报文	网络中的一些主机或设备发生故障时，会向交换机发送源 MAC 地址或目的 MAC 地址为全零的报文，配置该功能可丢弃这些全零报文并上报告警，管理员可根据告警信息来定位故障设备。**S2700SI/27010SI 系列不支持**
MAC 地址刷新 ARP 功能	配置 MAC 地址刷新 ARP 功能后，如果 MAC 地址表项的出接口发生变化，会及时更新 ARP 表项。**S2700SI/27010SI/2700EI 系列不支持**
端口桥功能	配置端口桥功能后，端口将处理相同源 MAC 地址和相同目的 MAC 地址的报文，适用于交换机下挂无二层转发能力的设备，或交换机作为数据中心的接入设备的网络场景。**S2700SI/27010SI/2700EI 系列不支持**

表15-1和表15-2列举的MAC地址表特性在支持的S系列交换机中都有对应的参数缺省配置，如表15-3所示。

表15-3 MAC地址表特性参数缺省值

参数	缺省值
动态 MAC 地址表项的老化时间	300s
MAC 地址学习功能	Enable
接口学习 MAC 地址的优先级	0
端口安全功能	Disabled
端口安全 MAC 地址学习限制数	1
端口安全功能的保护动作	Restrict
MAC 地址漂移表项的老化时间	300s
丢弃全 0 非法 MAC 地址报文的功能	Disabled
设备收到全 0 非法 MAC 地址报文的告警功能	Disabled
端口桥功能	Disabled
MAC 刷新 ARP 功能	Disabled

15.2 MAC地址表配置与管理

本节要介绍静态MAC地址表项、黑洞MAC地址表项、动态MAC地址表项、禁止MAC地址学习功能和

限制MAC地址学习数量的具体配置方法。

15.2.1 配置三种MAC地址表项

为了防止一些关键设备（如各种服务器或上行设备）被非法用户恶意修改其 MAC地址表项，可将这些设的MAC地址配置为静态MAC地址表项，因为静态MAC地址表项优先于动态MAC地址表项，不易被非法修改。

为了防止无用MAC地址表项占用MAC地址表，同时为了防止黑客通过MAC地址攻击用户设备或网络，可将那些有着恶意历史的非信任 MAC 地址配置为黑洞 MAC 地址，使设备在收到目的MAC或源MAC地址为这些黑洞MAC地址的报文时，直接予以丢弃，不修改原有的MAC地址表项，也不增加新的MAC地址表项。

为了减轻手工配置静态MAC地址表项，华为S系列交换机缺省已使能了动态MAC地址表项学习功能。但为了避免 MAC 地址表项爆炸式增长，可合理配置动态 MAC 表项的老化时间，以便及时删除 MAC 地址表中的废弃 MAC 地址表项。老化时间越短，交换机对周边的网络变化越敏感，适合在网络拓扑变化比较频繁的环境；老化时间越长，越适合在网络拓扑比较稳定的环境。

以上静态MAC地址表项、黑洞MAC地址表项、动态MAC地址表项这三种MAC地址表项的配置方法如表15-4所示。

表15-4 MAC地址表项配置步骤

步骤	命令	说明
1	**system-view** 例如：<HUAWEI> **system-view**	进入系统视图
2	**mac-address static** *mac-address interface-type interface-number* **vlan** *vlan-id* 例如：[HUAWEI] **mac-address static** 0001-0002-0003 **gigabitethernet** 0/0/2 **vlan** 4	添加静态 MAC 地址表项，相当于 MAC 地址与接口和 VLAN ID 进行绑定。命令中的参数说明如下。 （1）*mac-address*：指定要绑定的 MAC 的地址，格式为 H-H-H，其中 H 为 1 至 4 位的十六进制数，但不可为广播 MAC 地址、组播 MAC 地址和全零 MAC 地址 （2）*interface-type interface-number*：指定要绑定的出接口，也就是通过这个接口可以访问到以上 MAC 地址所对应的主机或其他设备。但该接口必须先加入下面由 *vlan-id* 参数指定的 VLAN 中，否则无法成功配置 （3）*vlan-id*：配置出接口所属的 VLAN 编号，相当于指定了以上接口所属 VLAN 的 ID，取值范围为 1～4 094 的整数 静态 MAC 地址表项的优先级高于动态 MAC 地址表项，如果通过 MAC 地址自动学习功能创建的 MAC 地址表项与原来的静态 MAC 地址表项相冲突，则该报文会被丢弃 可用 **undo mac-address static** *mac-address interface-type interface-number* **vlan** *vlan-id* 命令删除指定的静态 MAC 地址表项
3	**mac-address blackhole** *mac-address* [**vlan** *vlan-id* | **vsi** *vsi-name*] 例如：[HUAWEI]**mac-address blackhole** 0011-0022-0033 **vlan** 5	添加黑洞 MAC 地址表项。命令中的参数说明如下。 （1）*mac-address*：指定黑洞 MAC 地址表项中的 MAC 地址 （2）*vlan-id*：二选一可选参数，指定以上黑洞 MAC 地址所属 VLAN 的 ID，取值范围为 1～4 094 的整数 （3）*vsi-name*：二选一可选参数，指定以上黑洞 MAC 地址所属 VSI 实例的名称，为 1～31 个字符，不支持空格，区分大小写 可用 **undo mac-address blackhole** [*mac-address*] [**vlan** *vlan-id* | **vsi** *vsi-name*]命令删除指定的黑洞 MAC 地址表项
4	**mac-address aging-time** *aging-time* 例如：[HUAWEI] **mac-address aging-time** 600	配置动态 MAC 表项的老化时间，取值范围是 0 和 10～1000000 的整数秒，0 表示动态 MAC 地址表项不老化 缺省情况下，动态 MAC 表项的老化时间为 300s，可用 **undo mac-address aging-time** 命令恢复动态 MAC 地址表项的老化时间为缺省值

注意

在 MAC 地址表已满的情况下，继续配置静态或黑洞 MAC 表，则系统的处理方 法如下。

（1）如果MAC地址表中存在对应MAC地址的动态MAC地址表项，则添加的静态或黑洞MAC地址表项时自动覆盖原来对应的动态MAC地址表项。

（2）如果MAC地址表中不存在对应MAC地址的动态MAC地址表项，将无法添加静态或黑洞MAC地址表项。

在删除静态MAC地址表项、动态MAC地址表项和黑洞MAC地址表项时，如果不指定接口参数，将删除全部接口下对应的MAC地址表项；如果不指定VLAN参数，将删除全部 VLAN 下对应的 MAC 地址表项。正常情况下，一个 MAC 地址只可能对应一个MAC 地址表项，但有时可能由于所连接的设备移动到其他接口上，或者重新划分了接口所属VLAN，则会在交换机上创建相同MAC地址，但出接口或所属VLAN不同的表项。

15.2.2 配置禁止MAC地址学习功能

使能MAC地址学习功能时，收到来自周边设备的以太网帧后会从中解析出源MAC地址，再结合接收该以太网帧的接口和该接口所属VLAN的VLAN ID，在MAC地址表中添加新的表项。这样，以后设备接在收到去往该目的 MAC 地址的以太网帧时，则直接查询MAC地址表就可以得到正确的发送接口，可以避免广播。

如果想提高网络的安全性，防止设备学习到非法的 MAC 地址，错误地修改 MAC地址表中的原MAC地址表项，可以选择关闭设备上指定接口或者指定VLAN中所有接口的MAC地址学习功能，这样设备将不再从这些接口上学习新的MAC地址。

可以在接口视图下配置，仅禁止指定接口的MAC地址学习功能，也可以在VLAN视图下配置，禁止指定VLAN下所有接口的MAC地址学习功能，具体的配置步骤如表15-5所示。

表15-5 禁止MAC地址学习功能的配置步骤

步骤	命令	说明
1	**system-view** 例如：<HUAWEI> **system-view**	进入系统视图
在具体接口视图下配置		
2	**interface** *interface-type interface-number* 例如：[HUAWEI]**interface gigabitethernet 0/0/2**	键入要禁止 MAC 地址学习功能的接口（必须是二层接口），进入接口视图
3	**mac-address learning disable** [**action** { **discard** | **forward** }] 例如：[HUAWEI-GigabitEthernet0/0/2] **mac-address learning disable action discard**	在接口上禁止 MAC 地址学习功能。命令中的选项说明如下。 （1）**action discard**：二选一可选项，指定在收到报文后，对报文的目的 MAC 地址进行匹配，当与 MAC 地址表中某个表项匹配时，则对该报文进行转发，否则丢弃该报文 （2）**action forward**：二选一可选项，指定在收到报文后，直接按照报文中的目的 MAC 地址进行转发 如果不指定以上关闭 MAC 地址学习后的动作，则采用缺省的 **forward** 动作，但都不会通过学习报文中的 MAC 地址来生成新的 MAC 地址表项了 缺省情况下，接口的 MAC 地址学习功能是使能的，可用 **undo mac-address learning disable** 命令打开 MAC 地址学习功能

（续表）

步骤	命令	说明
在 VLAN 视图下配置		
2	**vlan** *vlan-id* 例如：[HUAWEI] **vlan** 5	（可选）键入要禁止接口 MAC 地址学习功能的 VLAN，进入 VLAN 视图
3	**mac-address learning disable** 例如：[HUAWEI-vlan5] **mac-address learning disable**	（可选）在 VLAN 中所有接口上禁止 MAC 地址学习功能。但在 VLAN 视图下不支持丢弃或转发动作的选择，都是 **forward** 动作，毕竟是针对 VLAN 中所有接口进行配置的，全部丢弃的话影响太大 缺省情况下，VLAN 的 MAC 地址学习功能是使能的，可用 **undo mac-address learning disable** 命令打开MAC 地址学习功能

15.2.3 配置限制MAC地址学习数量

交换机上是使用内存来保存这些MAC地址表项的，而交换机的内存容量是有限的，当黑客伪造大量源MAC 地址给交换机发送报文时，交换机的 MAC 表空间资源就可能被消耗尽，这样后面再收到合法用户的报文也无法学习新的源 MAC 地址，也不能创建新的MAC地址表项了。

为了解决以上问题，可以基于接口或者VLAN来限制对一些频繁遭到攻击的接口或者VLAN限制接口可以学习的MAC地址数量，当超过限制数量时，源MAC地址为新MAC地址的报文继续被转发，但是MAC地址表项不记录。另外，还可以配置发送告警动作，上报网管，从而防止MAC地址攻击，提高网络安全性。具体的配置步骤如表15-6所示。

表15-6 限制MAC地址学习数量的配置步骤

步骤	命令	说明
1	**system-view** 例如：<HUAWEI> **system-view**	进入系统视图
在接口视图下配置		
2	**interface** *interface-type interface-number* 例如：[HUAWEI]**interface gigabitethernet 0/0/2**	键入要禁止 MAC 地址学习功能的接口（必须是二层接口），进入接口视图
3	**mac-limit maximum** *max-num* 例如：[HUAWEI-GigabitEthernet0/0/2]**mac-limit maximum** 500	限制以上接口的 MAC 地址学习数量，不同系列交换机的取值范围不同，如 S2700 系列为 0～1 024 的整数，S3700 系列为 0～16 384 的整数，S5700/6700 系列为 0～4 096 的整数，S7700/9300/9300/9700 系列为 0～32 767 的整数。0 表示不限制 MAC 地址学习数量 缺省情况下，不限制 MAC 地址学习数量，可用 **undo mac-limit** 命令取消配置 MAC 地址学习限制
4	**mac-limit alarm** { **disable** | **enable** } 例如：[HUAWEI-GigabitEthernet0/0/2]**mac-limit a larm enable**	配置以上接口当 MAC 地址数量达到限制后是否进行告警。命令中的选项说明如下。 （1）**disable**：二选一选项，指定当 MAC 地址表项数目达到限制后，系统不发送告警 （2）**enable**：二选一选项，指定当 MAC 地址表项数目达到限制后，系统发送告警 缺省情况下，对超过 MAC 地址学习数量限制的报文进行告警，可用 **undo mac-limit alarm** 命令取消发送告警功能

（续表）

步骤	命令	说明
在 VLAN 视图下配置		
2	**vlan** *vlan-id* 例如：[HUAWEI] **vlan** 5	（可选）键入要配置接口 MAC 地址学习功能的 VLAN，进入 VLAN 视图
3	**mac-limit maximum** *max-num* 例如：[HUAWEI-vlan5] **mac-limit maximum** 1024	（可选）限制以上 VLAN 中的 MAC 地址学习数量。其他说明参见前面在接口视图下配置的第 3 步 缺省情况下，不限制 VLAN 中的 MAC 地址学习数量，可用 **undo mac-limit maximum** 命令取消配置 MAC 地址学习限制
4	**mac-limit alarm** { **disable** | **enable** } 例如：[HUAWEI-vlan5] **mac-limit alarm enable**	（可选）配置以上 VLAN 中当 MAC 地址数量达到限制后是否进行告警。其他说明参见前面在接口视图下配置的第 4 步 缺省情况下，对超过 MAC 地址学习数量限制的报文进行告警，可用 **undo mac-limit alarm** 命令取消发送告警功能

15.2.4 MAC地址表配置管理

配置好以上MAC地址表项后，可以通过以下display任意视图命令查看相关配置。

（1）display mac-address：查看所有MAC表项信息。

（2）display mac-address static：查看静态MAC表项信息。

（3）display mac-address dynamic：查看动态MAC表项信息。

（4）display mac-address blackhole：查看黑洞MAC表项信息。

（5）display mac-address aging-time：查看动态MAC表项的老化时间。

（6）display mac-address summary：查看设备上各种类型MAC地址表项的汇总信息。

（7）display mac-address total-number：查看设备上MAC地址表项的数量。

（8）display mac-limit：查看MAC地址学习数量限制信息。

15.2.5 MAC表配置示例

如图15-1所示，用户主机PC1的MAC地址为0002-0002-0002，用户主机PC2的MAC地址为 0003-0003-0003。LSW 连接到 Switch 上属于VLAN 2的GE0/0/1接口上。Server服务器的MAC地址为0004-0004-0004，连接的GE0/0/2接口也属于VLAN2。现在防止MAC地址欺骗攻击。

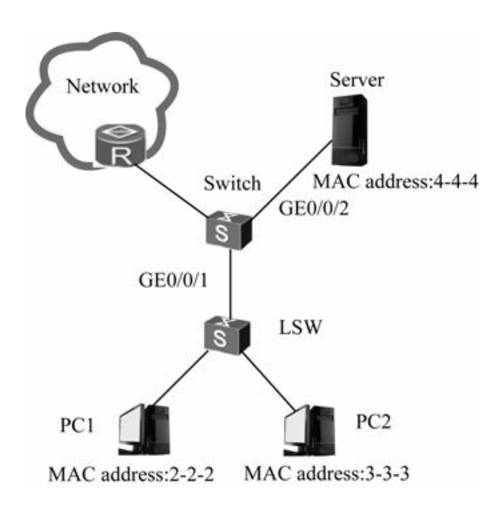


图15-1 MAC表配置示例拓扑结构

1．基本配置思路

为防止仿冒MAC地址的攻击，在Switch的MAC 表中为两个用户主机和服务器主机添加静态表项。同时为了进一步提高设备的安全性，还为动态MAC 地址表项设置一个相对较短的老化时间——500s。具体配置思路如下。

（1）创建所需的VLAN，并将接口加入对应的VLAN中，实现二层转发功能。

（2）添加静态MAC表项，实现防止MAC地址攻击。

（3）配置动态MAC表项的老化时间，以进一步提高网络的安全性。

2．具体配置步骤

（1）创建VLAN 2，配置GigabitEthernet0/0/1和GigabitEthernet0/0/2接口为不带VLAN标签的Hybrid类型（因为主机不能识别VLAN标签），然后加入VLAN 2中。

```
<Switch>system-view
[Switch] vlan 2
[Switch-vlan2] quit
[Switch] interface gigabitethernet 0/0/1
[Switch-GigabitEthernet0/0/1] port hybrid pvid vlan 2
[Switch-GigabitEthernet0/0/1] port hybrid untagged vlan 2
[Switch-GigabitEthernet0/0/1] quit
[Switch] interfacegigabitethernet 0/0/2
[Switch-GigabitEthernet0/0/2] port hybrid pvid vlan 2
[Switch-GigabitEthernet0/0/2] port hybrid untagged vlan 2
[Switch-GigabitEthernet0/0/2] quit
```

（2）为PC1、PC2和Server配置静态MAC地址表项，与它们对应的出接口和所属的VLAN的ID进行关联。

```
[Switch] mac-address static2-2-2GigabitEthernet0/0/1 vlan 2
[Switch] mac-address static3-3-3GigabitEthernet0/0/1 vlan 2
[Switch] mac-address static4-4-4GigabitEthernet0/0/2 vlan 2
```

（3）配置动态表项老化时间。

```
[Switch] mac-address aging-time500
```

配置好后，可在任意视图下执行display mac-address命令查看静态MAC表是否添加成功。具体如下，从中可以看到上面所配置的三条静态MAC地址表项。

```
[Switch] display mac-address static vlan 2
-------------------------------------------------------------------------------
MAC Address       VLAN/VSI       Learned-From       Type
-------------------------------------------------------------------------------
0002-0002-0002    2/-            GE0/0/1            static
0003-0003-0003    2/-            GE0/0/1            static
0004-0004-0004    2/-            GE0/0/2            static
-------------------------------------------------------------------------------
Total items displayed = 3
```

可在任意视图下执行display mac-address aging-time命令查看动态表项老化时间是否配置成功。具体如下，结果也是成功的，因为正确显示了老化时间为500s。

15.2.6 基于VLAN的MAC地址学习限制配置示例

如图15-2所示，用户网络1和用户网络2通过两台LSW与Switch相连，连接的接口分别为同属于VLAN 2的GigabitEthernet0/0/1和GigabitEthernet0/0/2接口。现为防止MAC地址欺骗攻击，控制接入用户数量，在Switch上配置对VLAN 2限制MAC地址学习功能。

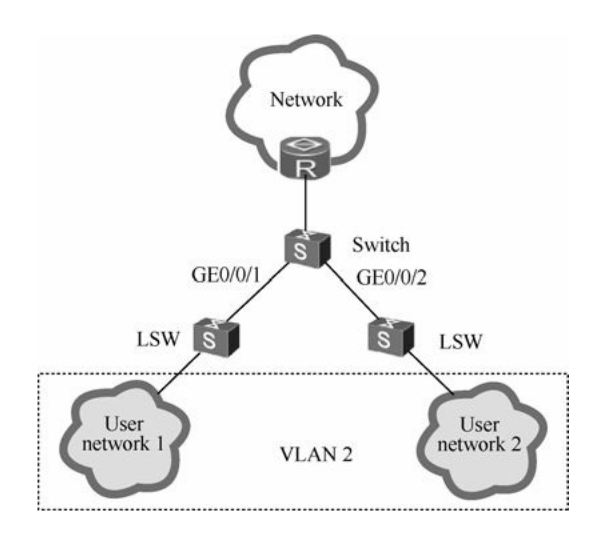


图15-2 基于VLAN的MAC地址学习限制的配置示例拓扑结构

1．基本配置思路

根据本示例要求可采用如下的思路配置基于VLAN的MAC地址学习限制。

（1）创建VLAN 2，并将对应接口加入VLAN 2中，实现二层转发功能。

（2）配置VLAN 2的MAC地址学习限制，实现防止MAC地址攻击，控制接入用户数量。

2．具体配置步骤

（1）与上节介绍的示例一样，创建VLAN 2，配置GigabitEthernet0/0/1和GigabitEthernet0/0/2接口为带VLAN标签的Hybrid类型（因为交换机间的连接必须带有VLAN标签）或者Trunk类型，然后加入VLAN 2中。

```
<Switch>system-view
[Switch] vlan 2
[Switch-vlan2] quit
[Switch] interface gigabitethernet 0/0/1
[Switch-GigabitEthernet0/0/1] port hybrid pvid vlan 2
[Switch-GigabitEthernet0/0/1] port hybrid tagged vlan 2
[Switch-GigabitEthernet0/0/1] quit
[Switch] interfacegigabitethernet 0/0/2
[Switch-GigabitEthernet0/0/2] port hybrid pvid vlan 2
[Switch-GigabitEthernet0/0/2] port hybrid tagged vlan 2
[Switch-GigabitEthernet0/0/2] quit
```

（2）在VLAN2上配置MAC地址学习限制规则：最多可以学习100个MAC地址，超过最大MAC地址学习数量的报文进行告警提示。

```
[Switch] vlan 2
[Switch-vlan2] mac-limit maximum100 alarm enable
[Switch-vlan2] quit
```

配置好后在任意视图下执行display mac-limit命令查看MAC地址学习限制规则。

15.3 端口安全配置与管理

端口安全（Port Security）功能是将设备端口学习到的MAC地址变为安全MAC地 址（包括安全动态MAC地址和Sticky MAC地址，是设备信任的MAC地址），以阻止除安全MAC和静态MAC之外的主机通过本接口和交换机通信，从而增强设备安全性。

在配置端口安全之前，需完成以下任务（因为它们与端口安全功能是相冲突的）：

（1）关闭基于端口的MAC地址学习限制功能。

（2）关闭配置的MUX VLAN功能。

（3）关闭MAC认证功能。

（4）关闭802.1x认证功能。

（5）关闭DHCP Snooping的MAC安全功能。

15.3.1 配置安全动态MAC功能

在对接入用户的安全性要求较高的网络中，可以配置端口安全功能，将接口学习到的MAC地址转换为安全动态MAC地址或Sticky MAC地址，且当接口上学习的最大MAC数量达到上限后不再学习新的MAC地址，只允许这些MAC地址和设备通信。这样可在一定程度上（因为非信任的MAC地址也可在达到最大可学习MAC地址数之前学习到）阻止其他非信任的MAC主机通过本接口和交换机通信，提高设备与网络的安全性。

缺省情况下，安全动态MAC表项不会被老化，但可以通过在接口上配置安全动态MAC 老化时间使其变为可以老化，且设备重启后安全动态 MAC 地址会丢失，需要重新学习 。安全动态MAC功能的配置步骤如表15-7所示。

表15-7 安全动态MAC功能的配置步骤

步骤	命令	说明
1	**system-view** 例如：<HUAWEI> **system-view**	进入系统视图
2	**interface** *interface-type interface-number* 例如：[HUAWEI]**interface gigabitethernet 0/0/2**	键入要配置安全动态 MAC 功能的接口（必须是二层接口），进入接口视图
3	**port-security enable** 例如：[HUAWEI-GigabitEthernet0/0/2] **port-security enable**	使能以上接口的端口安全功能。使能端口安全功能后，才可以配置端口安全保护动作、安全动态 MAC 学习限制数量和下节将要介绍的 Sticky MAC 功能 缺省情况下，未使能端口安全功能，可用 **undo port-security enable** 命令关闭该功能
4	**port-security max-mac-num** *max-number* 例如：[HUAWEI-GigabitEthernet0/0/2] **port-security max-mac-num** 100	（可选）配置以上接口的安全动态 MAC 学习限制数量，不同系列交换机的取值范围不同：S2700 系列为 1～1 024 的整数，S3700 系列为 1～16 384 的整数，S5700/6700/7700/9300/9300E/9700 系列为 1～4 096 的整数（其中 S5700LI、S5700S-LI、S5710EI 子系列为 1～1 024 的整数） 缺省情况下，接口学习的安全 MAC 地址限制数量为 1，可用 **undo port-security max-mac-num** 命令恢复端口安全 MAC 地址学习限制数为缺省值
5	**port-security protect-action { protect | restrict | shutdown }** 例如：[HUAWEI-GigabitEthernet0/0/2] **port-security protect-action protect**	（可选）配置以上接口的端口安全保护动作。命令中的选项说明如下。 （1）**protect**：多选一选项，指定当接口学习到的 MAC 地址数达到接口限制数时，丢弃源 MAC 地址不在 MAC 表中的报文

（续表）

步骤	命令	说明
5	**port-security protect-action { protect | restrict | shutdown }** 例如：[HUAWEI-GigabitEthernet0/0/2] **port-security protect-action protect**	（2）**restrict**：多选一选项，指定当接口学习到的 MAC 地址数超过接口限制数时，丢弃源 MAC 地址不在 MAC 表中的报文，并同时发出告警 （3）**shutdown**：多选一选项，指定当接口学习到的 MAC 地址数超过接口限制数时，将端口 error down（是一种管理关闭模式），同时发出告警。缺省情况下，端口关闭后不会自动恢复，只能由网络管理人员先执行 **shutdown** 命令再执行 **undo shutdown** 命令手动恢复，也可以在接口视图下执行 **restart** 命令重启接口 缺省情况下，端口安全保护动作为 **restrict**，可用 **undo port-security protect-action** 命令配置接口安全功能的保护动作为缺省动作
6	**port-security aging-time** *time* **[type { absolute | inactivity }]** 例如：[HUAWEI-GigabitEthernet0/0/2] **port-security aging-time** 30	（可选）配置以上接口学习到的安全动态 MAC 地址的老化时间。命令中的参数和选项说明如下。 （1）*time*：指定安全动态 MAC 地址的老化时间，取值范围为 1～1 440 的整数分钟 （2）**type absolute**：二选一可选项，配置安全动态 MAC 表项的老化类型为绝对时间老化，即系统每隔所设置的时间检测一次是否有该 MAC 地址的流量。如果没有流量，则立即将该安全动态 MAC 地址老化 （3）**type inactivity**：二选一可选项，配置安全动态 MAC 表项的老化类型为相对时间老化，即系统会每隔 1min 检测一次是否有该 MAC 地址的流量。如果没有流量，则经过所设置的时间后将该安全动态 MAC 地址老化 如果没有指定以上可选项，则缺省值为 **absolute**，即绝对时间老化类型 缺省情况下，接口学习的安全动态 MAC 地址不老化，可用 **undo port-security aging-time** 命令使该接口的安全动态 MAC 地址不老化

15.3.2 配置Sticky MAC功能

“Sticky（粘性） MAC地址”与上节介绍的“安全动态MAC地址”差不多，都属于安全 MAC 地址，都可在接口上使能端口安全功能后仅允许这些安全 MAC 地址和静态 MAC 地址与设备进行通信，在接口学习到的最大 MAC 数量达到上限后不再学习新的MAC地址。它们之间主要不同有以下几个方面。

（1）安全动态MAC地址可以通过在接口上配置老化时间来进行老化，但Sticky MAC地址永远不会被老化（不能通过在接口配置老化时间）。

（2）安全动态 MAC 地址对应的 MAC 表项在设备重启后丢失，需要重新学习，但Sticky MAC地址对应的MAC表项在设备重启后也不会丢失，无需重新学习。

（3）安全动态MAC地址表项只能通过动态学习得到，而Sticky MAC地址表项既可以通过安全动态MAC地址转换得到，又可以手工静态配置。

Sticky MAC功能特别适合为那些关键服务器或上行设备的MAC地址配置，因为永久有效，且所配置的Sticky MAC地址表项在设备重启后也不会丢失。Sticky MAC功能具体的配置步骤如表15-8所示，整体与上节介绍的安全动态MAC功能的配置差不多。

表15-8 Sticky MAC功能的配置步骤

步骤	命令	说明
1	**system-view** 例如：<HUAWEI> **system-view**	进入系统视图
2	**interface** *interface-type interface-number* 例如：[HUAWEI]**interface gigabitethernet 0/0/2**	键入要配置 Sticky MAC 功能的接口（必须是二层接口），进入接口视图
3	**port-security enable** 例如：[HUAWEI-GigabitEthernet0/0/2] **port-security enable**	使能以上接口的端口安全功能，其他说明参见上节表 15-8 中的第 3 步
4	**port-security mac-address sticky** 例如：[HUAWEI-GigabitEthernet0/0/2] **port-security mac-address sticky**	使能以上接口的 Sticky MAC 功能。使能 Sticky MAC 功能后接口会将学习到的动态 MAC 地址转化为 Sticky MAC（相当于静态 MAC） 缺省情况下，接口未使能 Sticky MAC 功能，可用 **undo port-security mac-address sticky** 命令去使能接口的 Sticky MAC 功能
5	**port-security max-mac-num** *max-number* 例如：[HUAWEI-GigabitEthernet0/0/2] **port-security max-mac-num** 100	（可选）配置以上接口安全动态 MAC 学习限制数量，其他说明参见上节表 15-8 中的第 4 步
6	**port-security protect-action** { **protect** | **restrict** | **shutdown** } 例如：[HUAWEI-GigabitEthernet0/0/2] **port-security protect-action protect**	（可选）配置以上接口的端口安全保护动作。其他说明参见上节表 15-8 中的第 5 步
7	**port-security mac-address sticky** *mac-address* **vlan** *vlan-id* 例如：[HUAWEI-GigabitEthernet0/0/2] **port-security mac-address sticky** 0001-0002-0003 **vlan** 5	（可选）手动配置 **sticky-mac** 地址表项。命令中的参数说明如下。 （1）*mac-address*：配置为 Sticky MAC 地址的 MAC 地址，格式为 H-H-H，其中 H 为 1 至 4 位的十六进制数，不能为 FFFF-FFFF-FFFF （2）*vlan-id*：指定以上 Sticky MAC 地址对应的出接口所属 VLAN 的 VLAN ID，取值范围为 1～4 094 缺省情况下，接口上没配置 Sticky MAC 地址，可用 **undo port-security mac-address sticky** [*mac-address* **vlan** *vlan-id*]命令删除指定的 Sticky MAC 地址

15.3.3 端口安全配置管理

配置好端口安全功能后，可使用以下display任意视图命令查看端口安全相关配置。

（1）display current-configuration interface interface-type interface-number：查看接口上的配置信息（包括端口安全配置信息）。

（2）display mac-address security [vlan vlan-id | interface-type interface-number] * [verbose]：查看指定VLAN或者指定接口，或者所有安全动态MAC表项的详细（选择verbose可选项时）或者摘信息。

（3）display mac-address sticky [vlanvlan-id | interface-type interface-number] * [verbose]：查看指定VLAN或者指定接口，或者所有Sticky MAC表项的详细（选择 verbose可选 项时）或者摘信息。

15.3.4 端口安全配置示例

本示例拓扑结构如图 15-3 所示，为了提高信息安全，将Switch连接用户侧的GE0/0/1接口使能端口安全功能，并且设置了端口学习MAC地址数的上限为信任的设备总数，这样其他外来人员使用自己带来的PC无法访问公司的网络。

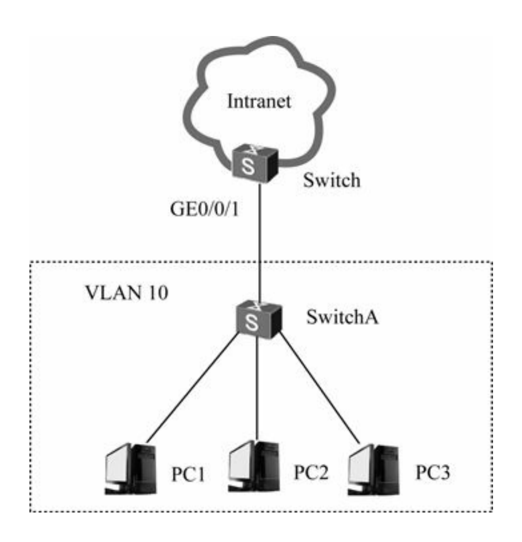


图15-3 端口安全配置示例拓扑结构

1．基本配置思路

本示例的要求很简单，就是要为接口使能端口安全功能。还可为了使设备在重启后不丢失所学习的安全功能，为这些信任设备使能 Sticky MAC功能，并配置当学习到的Sticky MAC地址超过限制的安全 MAC地址总数时为 protect 动作，丢弃源MAC地址不在MAC表中的报文。

2．具体配置步骤

（1）创建VLAN，配置接口类型，并把接口加到VLAN中。

```
<HUAWEI>system-view
[HUAWEI] sysname Switch
[Switch] vlan 10
[Switch-vlan10] quit
[Switch] interface gigabitethernet 0/0/1
[Switch-GigabitEthernet0/0/1] port link-type trunk
[Switch-GigabitEthernet0/0/1] port trunk allow-pass vlan 10
```

（2）配置GE0/0/1接口的端口安全功能和Sticky MAC功能，并配置安全功能动作。

```
[Switch-GigabitEthernet0/0/1] port-security enable    #---使能端口安全功能
[Switch-GigabitEthernet0/0/1] port-security mac-address sticky    #---使能接口Sticky MAC功能
[Switch-GigabitEthernet0/0/1] port-security protect-action protect#---配置端口安全功能的动作为“protect”动作
[Switch-GigabitEthernet0/0/1] port-security max-mac-num 4    #---配置接口学习MAC地址的数最多为4个
```

说明

这里之所以限制学习MAC地址数为4个，是假设Switch用户侧接了四台设备（三 台PC主机和接入交换机SwitchA），通过这样的限制，接口就不能再学习其他设备MAC地址了。这样如果PC1换成其他设备，就无法访问公司网络了。因为又配置的是Sticky MAC地址，即使SwitchA重启了，这些已学习到的MAC地址

表项也不会丢失，接口上连接成其他设备仍然不能访问网络，达到了示例中的要求。

15.4 其他基于MAC地址的安全功能配置

除了前面介绍两种主要的基于 MAC 地址的安全功能外，还有一些比较小的安全功能，如MAC地址防漂移功能、MAC地址漂移检测功能、MAC-spoofing-defend功能（禁止学习其他MAC地址）、丢弃全零MAC地址报文功能、MAC刷新ARP功能和端口桥功能等。本节对这些功能以及它们的具体配置方法分别予以介绍。

15.4.1 配置MAC地址防漂移

“MAC地址漂移”就是设备上一个接口学习到的MAC地址在同一VLAN中另一个接口上也被学习到，这样后面学习到的MAC地址表项就会覆盖原来的表项（对应的出接口不同了）。出现MAC地址漂移的原因主要有两个：（1）网络中交换机网线误接或配置错误形成了环网；（2）网络中某些非法用户仿冒合法的MAC地址进行MAC地址攻击。

配置MAC地址防漂移功能后，可以保证一个MAC地址的表项仅可在一个正确的接口上学习到，防止仿冒合法主机的MAC地址的入侵而改变该MAC地址原来正确的MAC地址表项。

MAC地址漂移功能的配置很简单，只需配置以下两项配置任务。

（1）配置接口MAC地址学习优先级

在接口上配置不同的 MAC 地址学习优先级后，如果不同接口学到相同的 MAC 地址表项，那么高优先级接口学到的 MAC 地址表项可以覆盖低优先级接口学到的 MAC地址表项，防止MAC地址发生漂移。

（2）配置不允许相同优先级接口MAC地址漂移

配置不允许相同优先级（缺省都是相同优先级）的接口发生 MAC 地址表项覆盖，也可以防止 MAC 地址漂移，提高网络的安全性。如设备的上行接口连接服务器，下行接口连接用户。为防止非法用户伪造服务器的 MAC 地址入侵，可以配置不允许相同优先级的接口发生 MAC 地址漂移。这样接口将不再学习相同的 MAC 地址，非法用户将无法使用网络设备MAC地址干扰设备与其他网络设备的正常通信。

但配置不允许相同优先级接口 MAC 地址漂移功能后也有负面的影响，如设备的接口连接的网络设备（例如：服务器）关机后，而设备的另外一个优先级相同的接口学习到与该网络设备同样的MAC地址（可能是伪造的），这时当原来关闭的网络设备再次上电后就不能再次正确学习这个设备的MAC地址，造成与该网络设备通信的中断。

以上两项MAC地址防漂移的配置任务的具体配置步骤如表15-9所示，都是可选的。

表15-9 MAC地址防漂移的配置步骤

步骤	命令	说明
1	**system-view** 例如：<HUAWEI> **system-view**	进入系统视图
2	**interface** *interface-type interface-number* 例如：[HUAWEI]**interface gigabitethernet 0/0/2**	键入要配置 MAC 地址防漂移功能的接口（必须是二层接口），进入接口视图
3	**mac-learning priority** *priority-id* 例如：[HUAWEI-GigabitEthernet0/0/2] **mac-learning priority** 2	配置接口学习 MAC 地址的优先级，取值范围为 0～3 的整数，数值越大优先级越高。**为想要正学习的某 MAC 地址的接口配置更高的优先级** 缺省情况下，所有接口学习 MAC 地址的优先级均为 0，可用 **undo mac-learning priority** 命令恢复为缺省值
4	**quit** 例如：[HUAWEI-GigabitEthernet0/0/2] **quit**	退出接口视图，返回系统视图

（续表）

步骤	命令	说明
5	**undo mac-learning priority** *priority-id* **allow-flapping** 例如：[HUAWEI] **undo mac-learning priority** 2 **allow-flapping**	全局配置不允许相同优先级的接口发生 MAC 地址漂移。参数 *priority-id* 用来指定不允许发生 MAC 地址漂移的接口学习 MAC 地址的优先级，取值范围为 0～3 的整数 缺省情况下，允许相同优先级的接口发生 MAC 地址漂移，可用 **mac-learning priority allow-flapping** 命令恢复缺省情况

15.4.2 MAC地址漂移检测配置与管理

上节介绍了设备的MAC地址防漂移功能，本节介绍的是MAC地址漂移检测功能。通过 MAC 地址漂移检测功能可以检测设备上所有的 MAC 地址是否发生了漂移，是一种早期预防功能。如果发生漂移，设备会上报告警到网管系统。

MAC地址漂移检测功能可以分别基于VLAN和全局配置，用户选择其一即可。

1．基于VLAN的MAC地址漂移检测

当基于 VLAN配置 MAC 地址漂移检测后，系统将检测该 VLAN内所有 MAC地址是否发生漂移，若出现 MAC 地址漂移则执行阻断动作，阻断时间到达后放开并重新进行检测。若 20s 内没有再次检测到 MAC 地址漂移，则接口阻塞被完全解除，重新开始一轮检测；若在20s内再次检测到MAC地址漂移，则再次开始阻塞，如此反复，直到达到设定的重试次数，若依然能够检测到 MAC 地址漂移，则永久阻断该接口。

当系统检测到某VLAN内有MAC地址发生漂移且发生漂移的接口或MAC地址被永久阻断时，只能通过配置解除指定VLAN下的接口阻断或MAC地址阻断来恢复到正常状态。

2．基于全局的MAC地址漂移检测

当基于配置全局 MAC 地址漂移检测功能，就可以检测到设备上所有的 MAC 地址是否发生了漂移。如果用户修改动态 MAC 表项的老化时间变长，会导致观测到 MAC地址漂移的时间变长，为了能够及时检测到 MAC 地址漂移，可以修改漂移表项的老化时间。当基于全局在端口上配置了 MAC 地址漂移处理动作后，如果系统检测到是该端口学习的MAC发生漂移，会将该端口关闭或者退出原来的VLAN。但在一个MAC地址漂移表项老化周期内只能关闭一个端口。

以上两种MAC地址漂移检测的具体配置方法如表15-10所示。

表15-10 MAC地址漂移检测的具体配置方法

步骤	命令	说明
1	**system-view** 例如：<HUAWEI> **system-view**	进入系统视图
基于 VLAN 配置 MAC 地址漂移检测功能		
2	**vlan** *vlan-id* 例如：[HUAWEI] **vlan** 10	键入要配置 MAC 地址漂移检测功能的 VLAN，进入 VLAN 视图

（续表）

步骤	命令	说明
3	**loop-detect eth-loop** { [**block-mac**] **block-time** *block-time* **retry-times** *retry-times* | **alarm-only** } 例如：[HUAWEI-vlan10] **loop-detect eth-loop block-mac** 0001-0002-0003 **retry-times** 3	配置对指定的 MAC 地址漂移检测功能。 （1）**block-mac**：可选项，指定根据 MAC 地址阻断。当没有指定此选项时，如果发现 MAC 地址漂移则阻断整个接口 （2）*block-time*：指定阻断的时间，取值范围为 10～65 535 的整数秒 （3）*retry-times*：指定阻断的重试次数，取值范围为 1～5 的整数 （4）**alarm-only**：二选一选项，指示当系统检测到 MAC 地址漂移时不阻断接口或 MAC 地址，只给网管发送告警 缺省情况下，没有配置基于 VLAN 的 MAC 地址漂移检测功能，可用 **undo loop-detect eth-loop** 命令取消对指定 VLAN 的该功能
4	**return** 例如：[HUAWEI-vlan10] **return**	退出 VLAN 视图，返回用户视图
5	**display loop-detect eth-loop** [**vlan** *vlan-id*] 例如：<HUAWEI> **display loop-detect eth-loop**	查看指定 VLAN 下 MAC 地址漂移的检测信息，包括被阻断的接口、所属 VLAN、被阻断的 MAC 地址等
基于全局配置 MAC 地址漂移检测功能		
2	**mac-address flapping detection** 例如：[HUAWEI] **mac-address flapping detection**	配置全局 MAC 地址漂移检测功能 缺省情况下，已经配置了全局 MAC 地址漂移检测功能，可用 **undo mac-address flapping detection** 命令取消配置全局 MAC 地址漂移检测功能
3	**mac-address flapping detection exclude vlan** { *vlan-id1* [**to** *vlan-id2*] } <1-10> 例如：[HUAWEI] **mac-address flapping detection exclude vlan** 5 **to** 10	（可选）配置 MAC 地址漂移检测的 VLAN 白名单，即指定不进行 MAC 地址漂移检测的 VLAN 缺省情况下，没有配置 MAC 地址漂移检测的 VLAN 白名单，可用 **undo mac-address flapping detection exclude vlan** { { *vlan-id1* [**to** *vlan-id2*] } <1-10> | **all** }命令删除 MAC 地址漂移检测时指定的或所有 VLAN 白名单
4	**mac-address flapping detection vlan** { { *vlan-id1* [**to** *vlan-id2*] } <1-10> | **all** } **security-level** { **high** | **middle** |**low** } 例如：[HUAWEI]**mac-address flapping detection vlan** 5 **security-level high**	（可选）配置指定 VLAN 中 MAC 地址漂移检测的安全级别。命令中的参数和选项说明如下。 （1）*vlan-id1* [**to** *vlan-id2*]：二选一参数，指定要配置 MAC 地址漂移检测的安全级别的 VLAN （2）<1-10>：表示 *vlan-id1* [**to** *vlan-id2*]参数可以最多有 10 个 （3）**all**：二选一选项，指定在所有 VLAN 上配置 MAC 地址漂移检测的安全级别 （4）**high**：多选一选项，配置对指定 VLAN 的 MAC 漂移检测安全级别为高，即 MAC 地址发生 3 次迁移后，系统认为发生了 MAC 地址漂移 （5）**middle**：多选一选项，配置对指定 VLAN 的 MAC 漂移检测安全级别为中，即 MAC 地址发生 10 次迁移后，系统认为发生了 MAC 地址漂移 （6）**low**：多选一选项，配置对指定 VLAN 的 MAC 漂移检测安全级别为高，即 MAC 地址发生 50 次迁移后，系统认为发生了 MAC 地址漂移 缺省情况下，MAC 地址漂移检测的安全级别为 **middle**，可用 **undo mac-address flapping detection vlan** { { *vlan-id1* [**to** *vlan-id2*] } <1-10> | **all** } **security-level** [**high** | **middle** | **low**]命令恢复指定 VLAN 的安全级别为缺省值

（续表）

步骤	命令	说明
5	**mac-address flapping aging-time** *aging-time* 例如：[HUAWEI] **mac-address flapping aging-time** 100	（可选）配置 MAC 地址漂移表项的老化时间，取值范围为 60～900 的整数秒 缺省情况下，MAC 地址漂移表项的老化时间为 300s，可用 **undo mac-address flapping aging-time** 命令恢复为缺省值
6	**interface** *interface-type interface-number* 例如：[HUAWEI]**interface gigabitethernet 0/0/1**	（可选）键入要配置发生 MAC 漂移后的处理动作的接口，进入接口视图
7	**mac-address flapping action** { **quit-vlan** | **error-down** } 例如：[HUAWEI-GigabitEthernet0/0/1] **mac-address flapping action quit-vlan**	（可选）配置接口发生 MAC 漂移后的处理动作。命令中的选项说明如下。 （1）**quit-vlan**：二选一选项，指定接口在发生 MAC 地址漂移后，该接口从原 VLAN 中退出 （2）**error-down**：二选一选项，指定接口在发生 MAC 地址漂移后，关闭该接口 缺省情况下，端口关闭后不会自动恢复，只能由网络管理人员先执行 **shutdown** 命令再执行 **undo shutdown** 命令手动恢复，也可以在接口视图下执行 **restart** 命令重启接口 缺省情况下，没有配置接口 MAC 地址漂移后的处理动作，可用 **undo mac-address flapping action** { **error-down** | **quit-vlan** }命令恢复缺省情况
8	**return** 例如：[HUAWEI-GigabitEthernet0/0/1] **return**	退出接口视图，返回用户视图
9	**display mac-address flapping** 例如：<HUAWEI> **display mac-address flapping**	（可选）查看 MAC 漂移检测功能的配置信息。可查看到的信息如下。 （1）是否已经配置了 MAC 地址漂移检测功能 （2）MAC 地址漂移表项的老化时间 （3）端口退 VLAN 的恢复时间 （4）MAC 地址漂移检测的 VLAN 白名单 （5）MAC 地址漂移检测 3 个安全级别的 VLAN 名单
10	**display mac-address flapping record** [**begin** *YYYY/MM/DD HH:MM:SS*] 例如：<HUAWEI> **display mac-address flapping record 2013/07/04 19:00:00**	（可选）查看 MAC 地址漂移的历史记录，可选参数 **begin** *YYYY/MM/DD HH:MM:SS* 用来指定要查看历史记录的开始时间。所能查到的信息同 **display mac-address flapping** 命令

15.4.3 配置MAC-spoofing-defend功能

这个“MAC-spoofing-defend”功能与前面介绍的MAC地址防漂移一样，最终的目的也是在一个接口学习到的 MAC 地址不允许再在其他接口上学习到，也可以防止某些非法用户仿冒MAC地址发送报文，只是所采用的方法不同而已。MAC-spoofing-defend功能是通过将接口配置为信任接口，以使该接口学习到的 MAC 地址在其他接口将不会再学习到。

MAC-spoofing-defend功能的配置方法很简单，只需要全局和对应接口上同时 使能该功能即可，具体配置步骤如表15-11所示。

表15-11 MAC-spoofing-defend功能的配置步骤

步骤	命令	说明
1	**system-view** 例如：<HUAWEI> **system-view**	进入系统视图
2	**mac-spoofing-defend enable** 例如：[HUAWEI] **mac-spoofing-defend enable**	使能全局 MAC-spoofing-defend 功能 缺省情况下，没有使能全局的 MAC-spoofing-defend 功能，可用 **undo mac-spoofing-defend enable** 命令去使能全局 MAC-spoofing-defend 功能
3	**interface** *interface-type interface-number* 例如：[HUAWEI]**interface gigabitethernet 0/0/2**	键入要配置 MAC 地址防漂移功能的接口（必须是二层接口），进入接口视图
4	**mac-spoofing-defend enable** 例如：[HUAWEI-GigabitEthernet0/0/2] **mac-spoofing-defend enable**	使能以上接口的 MAC-spoofing-defend 功能，即将接口配置为信任接口。如果信任接口连接的设备下电后，相应的动态 MAC 地址表项按照老化时间正常老化，**但如果更换信任接口连接的设备，则新设备的 MAC 地址不能被其他端口学习到** 缺省情况下，接口为不信任接口，可用 **undo mac-spoofing-defend enable** 命令取消接口为信任接口

15.4.4 配置丢弃全零MAC地址报文功能

网络中的一些主机或设备在发生故障时，往往会向交换机发送源 MAC 地址或目的MAC 地址为全 0 的报文。可配置交换机丢弃这些报文，还可以配置在收到这些报文时上报告警，管理员可根据告警信息来定位故障设备。具体的配置步骤如表15-12所示。

表15-12 丢弃全零MAC地址报文功能的配置步骤

步骤	命令	说明
1	**system-view** 例如：<HUAWEI> **system-view**	进入系统视图
2	**drop illegal-mac enable** 例如：[HUAWEI] **drop illegal-mac enable**	使能交换机丢弃全 0 非法 MAC 地址报文的功能 缺省情况下，交换机没有使能丢弃全 0 非法 MAC 地址报文的功能，可用 **undo drop illegal-mac enable** 命令恢复缺省情况
3	**drop illegal-mac alarm** 例如：[HUAWEI]**drop illegal-mac alarm**	（可选）配置交换机收到全 0 非法 MAC 地址报文时生成一条告警，**但只能告警一次**，如果需要继续告警，必须重新配置该命令。但在此之前一定要在设备系统视图下使用 **snmp-agent trap enable feature-name lldptrap** 命令使能设备的 LLDP 告警功能 缺省情况下，交换机收到全 0 非法 MAC 地址报文时不生成告警，可用 **undo drop illegal-mac alarm** 命令恢复缺省情况

15.4.5 配置MAC刷新ARP功能

如果用户更换某主机位置，使主机连接到设备的另一个接口上，该主机的MAC 地址将在另一个接口上被学习到，其对应的MAC表项的出接口将会变化，但是原ARP表项在到达老化时间后才会更新表项中的出接口。在到达老化时间之前，设备将使用错误的ARP表项来进行通信。配置MAC刷新ARP功能后，当MAC表项的出接口变化时可即时更新ARP表项。

MAC刷新ARP功能的配置方法很简单，只需在系统视图下使用mac-address update arp命令使能MAC刷新ARP功能即可。缺省情况下，没有使能MAC刷新ARP功能，可用undo mac-address update arp命令去使能MAC刷新ARP功能。

该命令只对动态ARP表项生效，不会更新静态ARP表项。另外，使用arp anti-attack entry-check { fixed-mac | fixed-all | send-ack } enable命令配置ARP表项固化功能（参见本书第16章的16.3.1节）后，MAC刷新ARP功能不生效。

15.4.6 配置端口桥功能

缺省情况下，设备在收到同源同宿报文（即源 MAC 地址和目的 MAC 地址所对应的MAC地址表项中的出接口均为设备上的同一接口的报文）时，设备判断为非法报文，并直接丢弃该报文。但有时，又的确存在源MAC地址和目的MAC地址所对应的MAC地址表项中的出接口为同一接口的情况，如下面两种情形。

（1）设备下挂有不具备二层转发能力的设备（如集线器）。当这些无二层转发能力的设备连接的用户有互通需求时，会直接将报文上送到上层二层设备，由二层设备完成转发功能。此时报文中的源MAC地址和目的MAC地址都会在二层设备的同一端口学习到。

（2）设备下连接一台启用了多个虚拟机的服务器。如果采用在服务器内部完成虚拟机之间的数据交换，会大大影响数据交换速度和服务器性能。为了提高数据交换速度和服务器性能，可由二层设备来进行数据交换。这样在设备的同一端口也会同时学习到报文中的源地址和目的MAC地址。

配置端口桥功能后，当端口收到同源同宿报文时，如果设备上的MAC 地址表中存在与该报文的目的MAC 地址对应的表项，则将报文从本端口转发出去。端口桥功能的配置也很简单，就是在对应的接口视图下使用port bridge enable命令使能端口桥功能即可。缺省情况下，没有配置端口桥功能，可用undoport bridge enable命令去使能对应接口的端口桥功能。

15.4.7 MAC防漂移配置示例

本示例拓扑结构如图15-4所示，某企业网络中的用户需要访问企业的服务器（Server）。由于在企业网中很难控制接入用户的行为，如果某些非法用户从其他接口假冒服务器的MAC地址发送报文，则服务器的MAC地址将在其他接口学习到。为了提高服务器安全性，防止被非法用户攻击，可在服务器连接的GE0/0/1 接口上配置MAC防漂移功能。

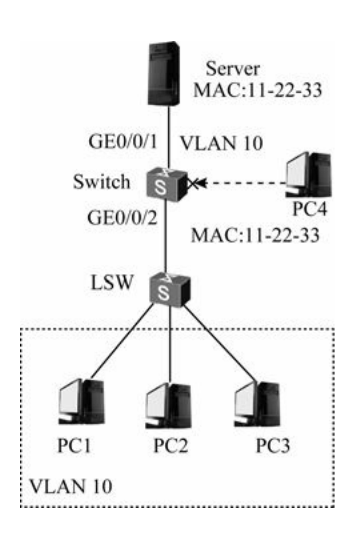


图15-4 MAC防漂移配置示例拓扑结构

因为本示例的要求和配置都非常简单，所以下面直接给出具体的配置步骤。

（1）创建VLAN，配置接口类型，并将接口加入VLAN中。

```
<Switch>system-view
[Switch] vlan 10
[Switch–vlan10] quit
[Switch] interface gigabitethernet 0/0/2
```

[Switch-GigabitEthernet0/0/2] port link-type trunk

[Switch-GigabitEthernet0/0/2] port trunk allow-pass vlan 10

[Switch-GigabitEthernet0/0/2] quit

[Switch] interface gigabitethernet 0/0/1

[Switch-GigabitEthernet0/0/1] port hybrid pvid vlan 10

[Switch-GigabitEthernet0/0/1] port hybrid untagged vlan 10

（2）在GigabitEthernet0/0/1接口上配置MAC地址学习的优先级为2，高于其他接口所采用的缺省值0，可以防止非法覆盖服务器的MAC地址表项。

[Switch-GigabitEthernet0/0/1] mac-learning priority 2

[Switch-GigabitEthernet0/0/1] quit

配置好后在任意视图下执行 display current-configuration 命令可查看接口MAC地址学习的优先级配置是否正确。

15.4.8 MAC地址漂移检测配置示例

本示例拓扑结构如图 15-5 所示，网络中两台 LSW 间网线可能因误接形成了网络环路，这样会引起MAC地址发生漂移、MAC地址表振荡。为了能够及时检测网络中出现的环路，可以在Switch上配置MAC地址漂移检测功能。

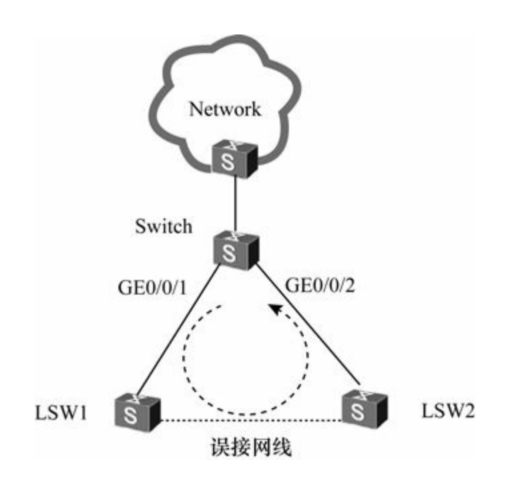


图15-5 MAC地址漂移检测配置示例拓扑结构

（1）开启MAC地址漂移检测功能。

<Switch>system-view

[Switch] mac-address flapping detection

（2）配置MAC地址漂移表项的老化时间为500s。

[Switch] mac-address flapping aging-time 500

（3）配置GE0/0/1和GE0/0/2接口MAC地址漂移后关闭。

[Switch] interface gigabitethernet 0/0/1

[Switch-GigabitEthernet0/0/1] mac-address flapping action error-down

[Switch-GigabitEthernet0/0/1] quit

[Switch] interface gigabitethernet 0/0/2

[Switch-GigabitEthernet0/0/2] mac-address flapping action error-down

[Switch-GigabitEthernet0/0/2] quit

（4）配置被Shutdown接口的自动恢复功能、自动恢复时间为500s。

[Switch] error-down auto-recovery cause mac-address-flapping interval500

配置完成后，当 GE0/0/1 接口的 MAC 地址漂移到接口 GE0/0/2 后，接口 GE0/0/2关闭；使用display mac-address flapping record命令可查看到漂移记录。

```
<Switch>display mac-address flapping record
S : start time
E : end time
(Q) : quit vlan
(D) : error down
-------------------------------------------------------------------------------
Move-Time              VLAN MAC-Address    Original-Port    Move-Ports    MoveNum
-------------------------------------------------------------------------------
S:2012-04-01 17:22:36    1   0000-0000-0007    GE0/0/1        GE0/0/2(D)    83
E:2012-04-01 17:22:44
-------------------------------------------------------------------------------
Total items on slot 0: 1
```

第16章　ARP安全配置与管理

说到ARP，这几年大家一定非常熟悉了，因为我们在日常的网络中经常提到“ARP病毒”、“ARP欺骗”、“ARP攻击”这几个新的术语，而且就是这小小的ARP协议经常令我们的网络莫名其妙地奇慢、上不了网，甚至导致设备或者整网络瘫痪，而且很难找到真正的原凶。有抓包经验的朋友记忆最深的可能就是在进行故障排除时，经常发现大量源IP地址或者MAC地址根本不是本地网络的广播包或帧，或者发现网关的MAC地址被非法地改了，甚至出现源或目的IP地址、MAC地址全为0的包，……。这一切都是因为ARP协议本身没有一个安全认证机制，给恶意攻击者留下了可乘之机。

总的来说，在ARP安全方面的隐患就是两个：一是利用非法、无效的伪造源/目的IP地址或者源MAC地址发送大量ARP报文，至于设备或者服务器无力承受而最终出现网络速度非常慢，或者中断，这就是通常所说的ARP泛洪攻击或者ARP DoS攻击；二是利用伪造的IP地址或者MAC地址非法修改网关或者网络中其他主机的ARP表项，造成用户无法上网，或者网络中无法正确访问这些主机，这就是通常所说的ARP网关或主机欺骗攻击。本章要全面介绍华为S系列交换机中针对以上两方面的ARP安全隐患而专门推出的许多具体ARP安全特性及配置与管理方法。

16.1 ARP安全概述

ARP（Address Resolution Protocol，地址解析协议）是将 IP地址解析为以太网MAC地址（或称物理地址）的协议。ARP协议有简单、易用的优点，但是也因为其没有任何安全认证机制而容易被攻击者利用。目前ARP攻击和ARP病毒已经成为局域网安全的一大威胁，为了避免各种攻击带来的危害，华为S系列交换机提供了多种技术对攻击进行检测和解决，这就是本章要介绍的主题。

ARP安全是针对ARP攻击的一种安全特性。它通过一系列对ARP表项学习和ARP报文处理的限制、检查等措施来保证网络设备的安全性。ARP安全特性不仅能够防范针对ARP协议的攻击，还可以防范网段扫描攻击等基于ARP协议的攻击。常见的ARP攻击方式主要包括以下两种。

（1）ARP泛洪攻击，也叫DoS（Denial of Service，拒绝服务）攻击，主要采用以下两种攻击方式。

① 攻击者通过伪造大量源**IP**地址变化 的ARP报文（以广播方式发送），使得设备ARP映射表缓存资源被无效的ARP表项耗尽（因为设备在接收到ARP报文后会提取报文中的源IP地址和源MAC地址，如果设备上没有对应的ARP映射表项就会生成新的ARP映射表项），造成合法用户的ARP报文不能继续生成ARP表项，最终导致正常用户的通信被中断。

② 攻击者利用工具扫描本网段主机或者进行跨网段扫描时，会向设备发送大量目的 **IP**地址不能解析的 IP报文，导致设备触发大量ARP Miss（ARP表项丢失）消息，生成并下发大量临时ARP表项，然后还会广播大量ARP请求报文以对目的IP地址进行解析，从而造成CPU负荷过重，直到瘫痪。

（2）ARP欺骗攻击，是指攻击者通过发送伪造的ARP报文（可以是伪造的免费ARP报文，也可以是伪造的ARP请求报文或ARP应答报文），非法修改设备或网络内其他用户主机的ARP表项，造成用户或网络

的报文通信异常。

为了避免上述各种ARP攻击带来的危害，华为S系列交换机提供了全方位针对ARP攻击进行防范、检测和解决的安全特性（大多数是在网关设备上部署的），具体如表16-1所示。

表16-1 针对泛洪和欺骗攻击的ARP安全特性

攻击类型	安全特性	功能说明	部署位置
ARP 泛洪	ARP 报文限速	通过限制接口接收 ARP 报文的速率（可以基于 ARP 报文中的源 IP 地址或源 MAC 地址进行 ARP 报文限速，也可基于全局、VLAN 或者接口进行 ARP 报文限速配置），可以防止设备因处理大量 ARP 报文而导致 CPU 负荷过重而无法处理其他业务的现象发生	建议在网关设备上部署。但当接入设备上部署了 MFF（MAC 地址强制转发）特性时，可在接入设备上针对全局、VLAN 和接口上部署
	ARP 报文源抑制	配置 ARP 报文源（指 ARP 报文中的源 IP 地址或源 MAC 地址）抑制功能后，可使当设备在一段时间内收到某一源 IP 地址或者源 MAC 的 ARP 报文数目超过设定阈值时，不再处理超出阈值部分的 ARP 请求报文，可以防止设置因处理大量某一源 IP 地址或者源 MAC 的 ARP 报文而致 CPU 负荷过重而无法处理其他业务的现象发生	建议在网关设备上部署

（续表）

攻击类型	安全特性	功能说明	部署位置
ARP 泛洪	ARP Miss 消息限速	通过对发送 ARP Miss 消息进行限速，可以防止设备因收到大量目的 IP 不能解析的 IP 报文而短时间内触发大量 ARP Miss 消息导致 CPU 负荷过重而无法处理其他业务的现象发生	建议在网关设备上部署
	ARP Miss 消息源抑制	配置 ARP Miss 消息源（仅指 ARP Miss 消息中的源 IP 地址）抑制功能后，如果一个源 IP 地址在一定时间内不断触发 ARP Miss 消息，当其速率超过了设定的阈值后，设备就认为此 IP 地址在进行攻击。此时，对于前 16 个攻击源，设备将下发 ACL 规则，在后续的一段时间内把这个地址发出的 IP 报文丢弃；对于之后的攻击源，设备使用设定的阈值对 IP 报文进行抑制。这样可以防止设备因收到某个源 IP 发送的大量目的 IP 不能解析的 IP 报文，触发大量 ARP Miss 消息，导致 CPU 负荷过重而无法处理其他业务的现象发生	
	免费 ARP 报文主动丢弃	免费 ARP 报文是指源 IP 地址和目的 IP 地址均为发送设备 IP 地址的 ARP 报文，主要用来通知其他设备更新针对自己的 ARP 表项。使能免费 ARP 报文主动丢弃功能后，设备直接丢弃免费 ARP 报文，可以防止设备因处理大量免费 ARP 报文，导致 CPU 负荷过重而无法处理其他业务的现象发生	建议在网关设备上部署（**仅 S7700/9300/9300E/9700 系列交换机支持**）
	ARP 表项严格学习	使能 ARP 表项严格学习功能后，只有本设备主动发送的 ARP 请求报文的应答报文才能触发本设备学习 ARP，其他设备主动向本设备发送的 ARP 报文不能触发本设备学习 ARP。这可以防止设备的缓存空间被无效的 ARP 表项占满	建议在网关设备上部署
	ARP 表项限制	使能 ARP 表项限制功能后，设备接口只能学习到设定的最大动态 ARP 表项数目，这样可以防止当一个接口所接入的某一台用户主机发起 ARP 攻击时整个设备的 ARP 表资源都被耗尽	
ARP 欺骗	ARP 表项固化（也称"防止 ARP 地址欺骗"）	使能 ARP 表项固化功能后，设备在第一次学习到 ARP 表项后不再允许用户更新此 ARP 表项，或只能更新此 ARP 表项的部分信息，或者需要通过发送 ARP 请求报文的方式进行确认，以防止攻击者伪造 ARP 报文修改正常用户的 ARP 表项内容 设备提供三种模式防御 ARP 地址欺骗攻击：fixed-all 模式、fixed-mac 模式和 send-ack 模式	
	动态 ARP 检测（也称"防止 ARP 中间人攻击"）	使能动态 ARP 检测（Dynamic ARP Inspection，DAI）功能后，当设备收到 ARP 报文时，将此 ARP 报文的源 IP 地址、源 MAC 地址、收到 ARP 报文的接口及 VLAN 信息和 DHCP Snooping 绑定表的信息进行比较，如果信息匹配则认为是合法用户，允许此用户的 ARP 报文通过，否则认为是攻击，丢弃该 ARP 报文 本功能仅适用于 DHCP Snooping 场景	建议在接入设备上部署。但当网关设备上部署了 DHCP 触发 ARP 学习特性时，则需要在网关设备上部署

（续表）

攻击类型	安全特性	功能说明	部署位置
ARP 欺骗	ARP 防网关冲突	通过 ARP 防网关冲突功能，可以防止用户仿冒网关发送 ARP 报文，非法修改网络内其他用户的 ARP 表项	建议在网关设备上部署
	免费 ARP 报文主动丢弃	使能免费 ARP 报文主动丢弃功能后，设备直接丢弃免费 ARP 报文，可以防止非法修改网关或用户主机的 ARP 表项	建议在网关设备上部署（仅 S7700/9300/9300E/9700 系列交换机支持）
	发送免费 ARP 报文	使能发送免费 ARP 报文功能后，网关设备主动向用户发送以自己 IP 地址为目的 IP 地址的 ARP 请求报文，定时更新用户 ARP 表项的网关 MAC 地址，防止用户的报文不能正常地转发到网关或者被恶意攻击者窃听	建议在网关设备上部署
	ARP 报文内 MAC 地址一致性检查	通过 ARP 报文内 MAC 地址一致性检查功能，可以防止以太网数据帧头部中的源/目的 MAC 地址和 ARP 报文数据区中的源/目的 MAC 地址不一致的 ARP 欺骗攻击	
	ARP 报文合法性检查	使能 ARP 报文合法性检查功能后，设备会对 MAC 地址和 IP 地址不合法的报文进行过滤。设备提供 3 种检查模式：源 MAC 地址、目的 MAC 地址和 IP 地址检查模式	建议在网关设备或接入设备上部署
	ARP 表项严格学习	使能 ARP 表项严格学习功能后，只有本设备主动发送的 ARP 请求报文的应答报文才能触发本设备学习 ARP，其他设备主动向本设备发送的 ARP 报文不能触发本设备学习 ARP。这可以防止设备因收到伪造的 ARP 报文，错误地更新 ARP 表项，导致合法用户的通信流量发生中断	建议在网关设备上部署
	DHCP 触发 ARP 学习	使能 DHCP 触发 ARP 学习功能后，设备根据收到的 DHCP ACK 报文直接生成 ARP 表项。当 DHCP 用户数目很大时，可以避免大规模 ARP 表项的学习和老化对设备性能和网络环境形成的冲击 此时设备上还可同时部署动态 ARP 检测功能，防止 DHCP 用户的 ARP 表项被伪造的 ARP 报文恶意修改	

16.2 配置防ARP泛洪攻击

ARP 泛洪攻击的基本思想就是发送大量的 ARP 报文，这样一方面可以使设备中用于缓存ARP表的内存资源被无效ARP表项耗尽，另一方面可能会使设备的CPU负荷过重造成用户无法正常通信。攻击者可能用来进行 ARP 泛洪攻击的报文包括 ARP 请求/应答报文、ARP Miss消息、免费ARP报文。表 16-2分析了不同报文类型的攻击方式和可用的ARP安全特性解决方案。当然，并不是所有情况下都需要配置全部的ARP安全特性，可根据实际选择其中一种或多种进行配置。

表16-2 ARP泛洪攻击方式及解决方案

报文类型	攻击方式	解决方案
ARP 请求应答报文	攻击者发送大量的 ARP 请求应答报文，造成设备 CPU 资源占用过高和 ARP 表项溢出	可采用的具体措施包括以下几种。 （1）ARP 报文限速：可基于源 MAC 地址、源 IP 地址对 ARP 报文进行限速，也可基于全局、VLAN 或接口进行 ARP 报文限速。**具体见 16.2.1、16.2.2 和 16.2.3 节** （2）ARP 表项限制：限制接口能够学习到的最大动态 ARP 表项数目，抑制攻击端口接收的 ARP 请求/应答报文时更新 ARP 表项。**具体见 16.2.8 节** （3）ARP 表项严格学习：使只有本设备主动发送的 ARP 请求报文的应答报文才能触发本设备学习 ARP 表项，其他设备主动向本设备发送的 ARP 报文不能触发本设备学习 ARP 表项，这样可以拒绝大部分的 ARP 报文攻击。**具体见 16.2.7 节**
ARP Miss 消息	攻击者使用大量找不到目的 MAC 地址的扫描 IP 报文攻击设备，造成设备产生大量的 ARP Miss 消息，生成 ARP 临时表项	可采用的具体措施包括以下几种。 （1）ARP Miss 消息源抑制：可基于源 MAC 地址、源 IP 地址对 ARP Miss 消息进行抑制。**具体见 16.2.4 节** （2）ARP Miss 消息限速：基于全局、VLAN 或者接口进行 ARP Miss 消息限速。**具体见 16.2.5 节** （3）配置 ARP 临时表项老化时间：使超时的 ARP Miss 报文设备不再处理。**具体见 16.2.6 节**
免费 ARP 报文	免费 ARP 是用于帮助作为网关的设备及时向 VLAN 内的主机更新 MAC 地址信息，防止对主机网关 MAC 地址的恶意篡改。但是，如果攻击者伪造大量的免费 ARP 报文可能造成设备 CPU 资源占用过高	可采取的具体措施为主动丢弃免费 ARP 报文（**仅 S7700/9300/9300E/9700 系列交换机支持**）。**具体见 16.2.9 节**

说明

当基于全局、VLAN、接口的ARP报文限速以及基于源MAC地址、源IP地址进 行 ARP 报文限速中的多个限速功能同时配置时，设备对同时满足这些限速条件的 ARP报文以其中最小的限速值进行限速。

当基于全局、VLAN、接口的ARP Miss消息限速以及基于源 IP地址进行ARP Miss消息限速中的多个限速功能同时配置时，设备对同时满足这些限速条件的ARP Miss消息以其中最小的限速值进行限速。

16.2.1 配置基于源**MAC**地址的**ARP**报文限速

设备处理大量源MAC地址相对固定的ARP报文会造成CPU繁忙，并且如果ARP报文的源IP地址同时不断变化，还会导致设备的ARP表资源被耗尽。为了避免此问题，可以配置设备根据源MAC地址进行ARP报文限速。设备会对上送CPU的ARP报文根据源MAC地址进行统计，如果在1s内收到的同一个源MAC地址的ARP报文超过设定阈值（ARP报文限速值），设备则丢弃超出阈值部分的ARP报文。

说明

本项 ARP 安全特性，在 S2700 系列中的 S2700SI 子系列和除 S2700-52P-EI、 S2700-52P-PWR-EI以外的S2700EI子系列均不支持；在S5700系列中，除S5700HI和S5710EI子系列外的其他型号均不支持；其他系列均支持。

基于源MAC地址的ARP报文限速的具体配置步骤如表16-3所示。

表16-3 基于源MAC地址的ARP报文限速配置步骤

步骤	命令	说明
1	**system-view** 例如：<HUAWEI> **system-view**	进入系统视图
2	**arp speed-limit source-mac maximum** *maximum* 例如：[HUAWEI] **arp speed-limit source-mac maximum** 100	配置**针对所有源 MAC 地址**的 ARP 报文源抑制速率。参数 *maximum* 用来限定基于源 MAC 地址的 ARP 报文速率，单位是 pps（每秒多少个报文）。在取值范围上，不同系列不太一样：S3700 系列的取值范围为 0～1 024 的整数，其他系列为 0～16 384 的整数。**如果取值为 0，表示不进行 ARP 报文源抑制** 缺省情况下，设备对每一个源 MAC 地址的 ARP 报文速率限制为 0，即不根据源 MAC 地址进行 ARP 报文限速，可用 **undo arp speed-limit source-mac** 命令将根据源 MAC 地址进行 ARP 限速的配置恢复为缺省配置
3	**arp speed-limit source-mac** *mac_addr* **maximum** *maximum* 例如：[HUAWEI] **arp speed-limit source-mac** 0-0-1 **maximum 50**	（可选）配置**针对指定源 MAC 地址**的 ARP 报文源抑制速率。命令中的参数说明如下。 （1）*mac_addr*：指定要进行 ARP 报文限速的源 MAC 地址，格式为 H-H-H，其中 H 为 4 位的十六进制数。取值为所有合法的单播 MAC 地址 （2）*maximum*：指定对应源 MAC 地址的 ARP 报文限制的速率，与上一步的该参数说明一样 **【说明】**对指定了源 MAC 地址的 ARP 报文源抑制速率为本步骤配置的 *maximum* 值；其他源 MAC 地址的 ARP 报文源抑制速率为步骤 2 中配置的 *maximum* 值 缺省情况下，所有 MAC 地址的 ARP 报文源抑制速率为 0pps，即不对 ARP 报文进行源抑制，可用 **undo arp speed-limit source-mac** *mac_addr* 命令将针对指定源 MAC 地址的 ARP 限速配置恢复为缺省配置

16.2.2 配置基于源IP地址ARP报文限速

与上节类似，但本节介绍的ARP报文限速是针对源**IP**地址相对固定（上节中介绍的是源**MAC**地址相对固定） 的ARP报文，会造成CPU繁忙，影响到正常业务的处理。为了避免此问题，可以配置设备根据源IP地址进行ARP报文限速。设备会对上送CPU的ARP报文根据源IP地址进行统计，如果在1s内收到的同一个源IP地址的ARP报文超过设定阈值，则丢弃超出阈值部分的ARP报文。

说明

本项 ARP 安全特性，在 S2700 系列中的 S2700SI 子系列和除 S2700-52P-EI、S2700-52P-PWR-EI以外的S2700EI子系列均不支持；在S5700系列中，除S5700HI和S5710EI子系列外的其他型号均不支持；其他系列均支持。

基于源IP地址的ARP报文限速的具体配置步骤如表16-4所示。它与上节介绍的基于源MAC地址的ARP报文限速配置基本一样。

表16-4 基于源IP地址的ARP报文限速配置步骤

步骤	命令	说明
1	**system-view** 例如：<HUAWEI> **system-view**	进入系统视图
2	**arp speed-limit source-ip maximum** *maximum* 例如：[HUAWEI] **arp speed-limit source-ip maximum** 100	配置针对**所有源 IP 地址**的 ARP 报文源抑制速率，参数 *maximum* 用来限定基于源 IP 地址的 ARP 报文速率，单位是 pps。S3700 系列的取值范围为 0～1 024，其他系列为 0～16 384。如果取值为 0，表示不进行 ARP 报文源抑制 缺省情况下，S2700/3700/5700/6700 系列交换机对每一个源 IP 地址的 ARP 报文速率限制为 0，即不根据源 MAC 地址进行 ARP 报文限速；而 S7700/9300/9300E/9700 系列交换机允许 1s 内最多只能有 30 个同一个源 IP 地址的 ARP 报文通过，可用 **undo arp speed-limit source-ip** 命令将根据源 IP 地址进行 ARP 限速的配置恢复为缺省配置
3	**arp speed-limit source-ip** *ip-address* **maximum** *maximum* 例如：[HUAWEI] **arp speed-limit source-ip** 192.168.10.1 **maximum 50**	（可选）配置**针对指定 IP 地址**的 ARP 报文源抑制速率。命令中的参数说明如下。 （1）*ip-address*：指定要进行 ARP 报文限速的源 IP 地址，点分十进制格式 （2）*maximum*：指定对应源 IP 地址的 ARP 报文限制的速率，与上一步的该参数说明一样 【说明】对指定了源 IP 地址的 ARP 报文源抑制速率为本步骤配置的 *maximum* 值；其他源 IP 地址的 ARP 报文源抑制速率为步骤 2 中配置的 *maximum* 值 缺省情况下，S2700/3700/5700/6700 系列交换机对每一个源 IP 地址的 ARP 报文速率限制为 0，即不根据源 MAC 地址进行 ARP 报文限速，而 S7700/9300/9300E/9700 系列交换机允许 1s 内最多只能有 30 个同一个源 IP 地址的 ARP 报文通过，可用 **undo arp speed-limit source-ip** *ip-address* 命令将针对指定源 IP 地址的 ARP 限速配置恢复为缺省配置

16.2.3 配置基于全局、VLAN或者接口的ARP报文限速

在ARP泛洪攻击的报文中，如果源MAC地址，或者源IP地址都不是相对固定的情况下，可以在全局、VLAN 或接口下配置针对所有 ARP 报文的限速和限速时间。在ARP报文限速时间内，如果收到的所有ARP报文数目超过ARP报文限速值，设备会丢弃超出限速值的ARP报文。

（1）全局的ARP报文限速：在设备出现ARP攻击时，需要限制全局处理的ARP报文数量（是指设备各个接口上接收到的ARP报文的总数）。

（2）VLAN的ARP报文限速：在某个VLAN内的所有接口出现ARP攻击时，可以仅限制处理收到的该VLAN内的ARP报文数量，配置本功能可以保证不影响其他VLAN内所有接口的ARP学习。

（3）接口的ARP报文限速：在某个接口出现ARP攻击时，可以仅限制处理该接口（可以是各种以太网接口，也可以是**Eth-Trunk**接口，还可以是端口组 ）收到的ARP报文数量，配置本功能可以保证不影响其他接口的ARP学习。

当同时在全局、VLAN或接口下配置ARP报文的限速值和限速时间时，设备会先按照接口进行限速，再按照VLAN进行限速，最后按照全局进行限速。当设备丢弃的ARP报文数量较多时，如果希望设备能够以告警的方式提醒网络管理员，则还可以使能ARP 报文限速丢弃告警功能。当丢弃的ARP报文数超过告警阈值时，设备将产生告警。

说明

本项 ARP 安全特性，在 S2700 系列中的 S2700SI 子系列和除 S2700-52P-EI、 S2700-52P-PWR-EI以外的S2700EI子系列，以及S5700LI和S5700S-LI子系列均不支持；其他系列均支持。

基于全局、VLAN或接口的ARP报文限速的具体配置步骤如表16-5所示。这几种针对不同范围的ARP报文限速配置之间的主要区别就是配置所在视图不一样，主要配置命令完全一样。

表16-5 基于全局、VLAN或者接口的ARP报文限速配置步骤

步骤	命令	说明
1	**system-view** 例如：<HUAWEI> **system-view**	进入系统视图
2	**interface** *interface-type interface-number* 例如：[HUAWEI] **interface gigabitethernet 1/0/0** 或：**vlan** *vlan-id* 例如：[HUAWEI] **vlan** 100	（可选）键入要配置 ARP 报文限速的接口，进入接口视图，或者键入要配置 ARP 报文限速的 VLAN，进入 VLAN 视图 【说明】*在系统视图下配置 ARP 报文限速功能无需执行此步骤*
3	**arp anti-attack rate-limit enable** 例如：[HUAWEI] **arp anti-attack rate-limit enable** 或：[HUAWEI-GigabitEthernet1/0/0] **arp anti-attack rate-limit enable** 或：[HUAWEI-Vlanif100] **arp anti-attack rate-limit enable**	在全局，或者 VLAN，或者接口上使能 ARP 报文限速功能。缺省情况下，没有使能 ARP 报文限速功能，可用 **undo arp anti-attack rate-limit enable** 命令去使能 ARP 报文速率抑制功能
4	**基于全局或 VLAN 配置时：** **arp anti-attack rate-limit** *packet-number* [*interval-value*] 例如：[HUAWEI] **arp anti-attack rate-limit** 200 20 或：[HUAWEI-Vlanif100] **arp anti-attack rate-limit** 200 20 **基于接口配置时：** **arp anti-attack rate-limit** *packet-number* [*interval-value* | **block timer** *timer*]* 例如：[HUAWEI-GigabitEthernet1/0/0] **arp anti-attack rate-limit** 200 20 **block timer** 60	在全局、VLAN 或接口下配置 ARP 报文的速率抑制功能，包括配置 ARP 报文的限速值和限速时间，以及当某个接口的 ARP 报文超过限速值时，在后续一段时间内持续丢弃该接口下收到的所有 ARP 报文的功能（即开启 **block** 模式）。命令中的参数说明如下。 （1）*packet-number*：指定 ARP 报文的限速值，即限速时间内允许通过的 ARP 报文的个数，取值范围 1～16 384 的整数，缺省值为 100 （2）*interval-value*：可多选参数，指定 ARP 报文的限速时间，取值范围 1～86 400 的整数秒，缺省值为 1s （3）**block timer** *timer*：可多选参数，**仅在接口下配置时选用**，用于指定持续丢弃超过 ARP 报文限速值的接口下收到的所有 ARP 报文的时长，取值范围 1～86 400 的整数秒，缺省值为 1s。**但系统视图和 VLAN 视图下不支持该参数，S2700/3700 系列交换机不支持该参数** 【说明】*该命令在非* **block** *模式下只对上送 CPU 的 ARP 报文进行限速，对芯片转发的报文不会产生影响；在* **block** *模式下，仅在接口下上送 CPU 的 ARP 报文超过限速值时会触发* **block**，*触发后设备会持续丢弃该接口下的所有 ARP 报文* 缺省情况下，在 1s 内最多允许 100 个 ARP 报文通过，且没有配置当某个接口的 ARP 报文超过限速值时在后续一段时间内持续丢弃该接口下收到的所有 ARP 报文的功能，可用 **undo arp anti-attack rate-limit** 命令将全局、VLAN 或接口下配置的 ARP 报文的限速值和限速时间恢复为缺省值，并恢复 ARP 报文的上送

（续表）

步骤	命令	说明
5	**arp anti-attack rate-limit alarm enable** 例如：[HUAWEI] **arp anti-attack rate-limit alarm enable** 或：[HUAWEI-GigabitEthernet1/0/0]**arp anti-attack rate-limit alarm enable** 或：[HUAWEI-Vlanif100] **arp anti-attack rate-limit alarm enable**	（可选）在全局、VLAN 或接口下使能 ARP 报文限速丢弃告警功能。这样，当丢弃的 ARP 报文数超过告警阈值时，设备将产生告警 缺省情况下，没有使能 ARP 报文限速丢弃告警功能，可用 **undo arp anti-attack rate-limit alarm enable** 命令去使能 ARP 报文限速丢弃告警功能
6	**arp anti-attack rate-limit alarm threshold** *threshold* 例如：[HUAWEI] **arp anti-attack rate-limit alarm threshold** 200 或：[HUAWEI-GigabitEthernet1/0/0]**arp anti-attack rate-limit alarm threshold** 100 或：[HUAWEI-Vlanif100]**arp anti-attack rate-limit alarm threshold** 100	（可选）在全局、VLAN 或接口下配置 ARP 报文限速丢弃告警阈值，取值范围为 1～16 384 的整数。通过本命令可以配置告警阈值，当设备丢弃因超过 ARP 限速值的 ARP 报文数超过告警阈值时，设备将以告警的方式通知网络管理员 缺省情况下，ARP 报文限速丢弃告警阈值为 100，可用 **undo arp anti-attack rate-limit alarm threshold** 命令恢复 ARP 报文限速丢弃告警阈值为缺省值

16.2.4 配置ARP Miss消息源抑制

当设备检测到某一源 IP地址的 IP报文在1s内触发的ARP Miss消息数量超过了限速值，就认为此源 IP地址存在攻击。这时，设备对ARP Miss报文的缺省处理方式是block方式，即设备会丢弃超出限速值部分的ARP Miss消息，也就是丢弃触发这些ARP Miss消息的ARP Miss报文，并下发一条ACL来丢弃该源IP地址的后续所有ARP Miss报文；如果是none-block方式，设备只会丢弃超出限速值部分的ARP Miss消息，因此，该

方式对CPU的负担减轻效果有限。可根据实际情况调整ARP Miss消息的限速值并合理配置ARP Miss报文处理方式。

说明

本项 ARP 安全特性，在 S2700 系列中的 S2700SI 子系列和除 S2700-52P-EI、 S2700-52P-PWR-EI以外的S2700EI子系列，以及S5700LI和S5700S-LI子系列均不支持；其他系列均支持。

ARP Miss消息源抑制功能的具体配置步骤如表 16-6所示。

表16-6 ARP Miss消息源抑制的配置步骤

步骤	命令	说明
1	**system-view** 例如：<HUAWEI> **system-view**	进入系统视图
2	**arp-miss speed-limit source-ip maximum** *maximum* 例如：[HUAWEI] **arp-miss speed-limit source-ip maximum** 100	配置针对所有 ARP Miss 消息的源抑制功能，参数 *maximum* 用来指定基于源 IP 地址的 ARP Miss 消息的抑制速率，取值范围为 0～16 384 的整数。如果取值为 0，表示不根据源 IP 地址进行 ARP Miss 消息限速 缺省情况下，ARP Miss 消息的源抑制功能已经使能，S7700/9300/9300E/9700 系列交换机允许每秒最多处理 30 个（S2700/3700/5700/6700 系列为 500 个）同一个源 IP 地址触发的 ARP Miss 消息，如果同一个源 IP 地址在 1s 内触发的 ARP Miss 消息个数超过 ARP Miss 消息限速值，设备会丢弃超过限速值的 ARP Miss 消息，可用 **undo arp-miss speed-limit source- ip** 命令将 ARP Miss 消息源抑制速率限制恢复为缺省配置

（续表）

步骤	命令	说明
3	**arp-miss speed-limit source-ip** *ip-address* [**mask** *mask*]**maximum** *maximum* [**none-block** | **block timer** *timer*] 例如：[HUAWEI] **arp-miss speed-limit source-ip** 192.168.10.1 **maximum** 50	（可选）配置针对指定源 IP 地址的 ARP Miss 消息源抑制功能，并指定 ARP Miss 报文处理方式。命令中的参数和选项说明如下。 （1）*ip-address*：指定 IP 地址，表示对特定 IP 地址用户的 ARP Miss 报文进行源速率抑制 （2）*mask*：可选参数，指定以上 IP 地址的子网掩码，针对该网段用户的 ARP Miss 消息进行限速 （3）*maximum*：指定基于源 IP 地址的 ARP Miss 消息限速值，取值范围为 0～16 384 的整数。如果取值为 0，表示不根据源 IP 地址进行 ARP Miss 消息限速 （4）**none-block**：二选一可选项，指定 ARP Miss 报文处理方式为 **none-block**。即指定源 IP 地址的 IP 报文在 1s 内触发的 ARP Miss 消息个数超过限速值时，设备只做软件限速（CPU 处理），丢弃超过限速值的 ARP Miss 消息，即丢弃触发这些 ARP Miss 消息的 ARP Miss 报文 （5）**block timer** *timer*：二选一可选参数，指定 ARP Miss 报文处理方式为 **block**。即一旦指定源 IP 地址的 IP 报文在 1s 内触发的 ARP Miss 消息个数超过限速值，设备会丢弃超过限速值的 ARP Miss 消息，即丢弃触发这些 ARP Miss 消息的 ARP Miss 报文，同时设备会下发一个 ACL 让芯片在 *timer* 时间内持续丢弃该源 IP 地址的后续所有 ARP Miss 报文。超过该时间后，ACL 将被老化，芯片不再丢弃报文，报文将恢复上送 CPU 处理 【说明】两种本步与上述第 2 步同时配置，则当触发 ARP Miss 消息的 IP 报文的源 IP 地址匹配本步限速指定的 IP 地址时，对该源 IP 地址的 IP 报文触发的 ARP Miss 消息限速值为本步配置值；否则为第 2 步中配置的 *maximum* 值 缺省情况下，ARP Miss 消息的源抑制功能已经使能，S7700/9300/9300E/9700 系列交换机允许每秒最多处理 30 个（S2700/3700/5700/6700 系列为 500 个）同一个源 IP 地址触发的 ARP Miss 消息。如果同一个源 IP 地址在 1s 内触发的 ARP Miss 消息个数超过 ARP Miss 消息限速值，设备会丢弃超过限速值的 ARP Miss 消息，并缺省使用 **block** 方式在 5s 内持续丢弃该源 IP 地址的后续所有 ARP Miss 报文，可用 **undo arp-miss speed-limit source-ip** [*ip-address* [**mask** *mask*]]命令将根据指定源 IP 地址进行 ARP Miss 消息限速的配置恢复为缺省配置

16.2.5 配置全局、VLAN和接口的ARP Miss消息限速

如果网络中有用户向设备发送大量目的IP地址不能解析的IP报文（即路由表中存在该IP报文的目的IP地址对应的路由表项，但设备上没有该路由表项中下一跳对应的ARP表项），将导致设备触发大量的ARP Miss消息。这种触发ARP Miss消息的 IP报文会被上送到主控板进行处理，设备会根据ARP Miss消息生成和下发大量临时ARP表项并向目的网络发送大量ARP请求报文，这样就增加了设备CPU的负担，同时严重消耗目的网络的带宽资源。

为了避免这种 IP报文攻击所带来的危害，可配置以下ARP Miss消息限速功能。

（1）全局的ARP Miss消息限速：在设备出现目的 IP地址不能解析的 IP报文攻击时，可限制全局处理的ARP Miss消息数量。

（2）VLAN的ARP Miss消息限速：在某个VLAN内的所有接口出现目的 IP地址不能解析的 IP报文攻击时，可仅限制处理该VLAN内报文触发的ARP Miss消息数量，配置本功能可以保证不影响其他VLAN内所有接口的IP报文转发。

（3）接口的ARP Miss消息限速：在某个接口出现目的 IP地址不能解析的 IP报文攻击时，可仅限制处理该接口收到的报文触发的ARP Miss消息数量，配置本功能可以保证不影响其他接口的IP报文转发。

当同时在全局、VLAN或接口下配置ARP Miss消息限速时，设备会先按照接口进行限速，再按照VLAN进行限速，最后按照全局进行限速。当设备忽略的ARP Miss消息数量较多时，如果希望设备能够以告警的方式提醒网络管理员，则可以配置ARP Miss消息限速丢弃告警功能。当设备忽略处理的ARP Miss消息个数超过告警阈值时，设备将产生告警。

说明

本项 ARP 安全特性，在 S2700 系列中的 S2700SI 子系列和除 S2700-52P-EI、S2700-52P-PWR-EI以外的S2700EI子系列，以及S5700LI和S5700S-LI子系列不均支持；其他系列均支持。

基于全局、VLAN或接口的ARP Miss消息限速的具体配置步骤如表 16-7所示。这几种针对不同范围的ARP Miss限速配置之间的主要区别就是配置所在视图不一样，主要配置命令完全一样。且与16.2.3节介绍的基于全局、VLAN或接口的ARP报文限速配置基本一样，只是命令名称上的少许差别。

表16-7 基于全局、VLAN或者接口的ARP Miss限速配置步骤

步骤	命令	说明
1	**system-view** 例如：<HUAWEI> **system-view**	进入系统视图
2	**interface** *interface-type interface-number* 例如：[HUAWEI] **interface gigabitethernet 1/0/0** 或：**vlan** *vlan-id* 例如：[HUAWEI] **vlan** 100	（可选）键入要配置 ARP Miss 消息限速的接口，进入接口视图，或者键入要配置 ARP Miss 消息限速的 VLAN，进入 VLAN 视图 【说明】在系统视图下配置 ARP Miss 消息限速功能无需执行此步骤
3	**arp-miss-miss anti-attack rate-limit enable** 例如：[HUAWEI] **arp-miss-miss anti-attack rate-limit enable** 或：[HUAWEI-GigabitEthernet1/0/0] **arp-miss-miss anti-attack rate-limit enable** 或：[HUAWEI-Vlanif100] **arp-miss-miss anti-attack rate-limit enable**	在全局，或者 VLAN，或者接口上使能 ARP Miss 消息限速功能。缺省情况下，没有使能 ARP Miss 消息限速功能，可用 **undo arp-miss anti-attack rate-limit enable** 命令去使能 ARP Miss 消息速率抑制功能
4	**arp-miss anti-attack rate-limit** *packet-number* [*interval-value*] 例如：[HUAWEI] **arp-miss anti-attack rate-limit** 200 20 或：[HUAWEI-GigabitEthernet1/0/0] **arp-miss anti-attack rate-limit** 200 20 **block timer** 60 或：[HUAWEI-Vlanif100] **arp-miss anti-attack rate-limit** 200 20	在全局、VLAN 或接口下配置 ARP Miss 消息的速率抑制功能，包括配置 ARP Miss 消息的限速值和限速时间。在 ARP Miss 消息限速时间内，如果收到的 IP 报文触发的 ARP Miss 消息数目超过 ARP Miss 消息限速值，设备将忽略处理超出限速值的 ARP Miss 消息，并丢弃超出限速值的触发 ARP Miss 消息的 IP 报文（即 ARP Miss 报文）。命令中的参数说明如下。

（续表）

步骤	命令	说明
4	**arp-miss anti-attack rate-limit** *packet-number* [*interval-value*] 例如：[HUAWEI] **arp-miss anti-attack rate-limit** 200 20 或：[HUAWEI-GigabitEthernet1/0/0] **arp-miss anti-attack rate-limit** 200 20 **block timer** 60 或：[HUAWEI-Vlanif100] **arp-miss anti-attack rate-limit** 200 20	（1）*packet-number*：指定 ARP Miss 消息的限速值，即限速时间内允许通过的 ARP Miss 消息的个数，取值范围 1～16 384 的整数，缺省值为 100 （2）*interval-value*：可多选参数，指定 ARP Miss 消息的限速时间，取值范围 1～86 400 的整数秒，缺省值为 1s 缺省情况下，在 1s 内最多允许 100 个 ARP Miss 消息通过，可用 **undo arp-miss anti-attack rate-limit** 命令将全局、VLAN 或接口下配置的 ARP Miss 消息的限速值和限速时间恢复为缺省值
5	**arp-miss anti-attack rate-limit alarm enable** 例如：[HUAWEI] **arp-miss anti-attack rate-limit alarm enable** 或：[HUAWEI-GigabitEthernet1/0/0] **arp-miss anti-attack rate-limit alarm enable** 或：[HUAWEI-Vlanif100] **arp-miss anti-attack rate-limit alarm enable**	（可选）在全局、VLAN 或接口下使能 ARP Miss 消息限速丢弃告警功能。当丢弃的 ARP Miss 消息数超过告警阈值时，设备将产生告警 缺省情况下，没有使能 ARP Miss 消息限速丢弃告警功能，可用 **undo arp-miss anti-attack rate-limit alarm enable** 命令去使能 ARP Miss 消息限速丢弃告警功能
6	**arp-miss anti-attack rate-limit alarm threshold** *threshold* 例如：[HUAWEI] **arp-miss anti-attack rate-limit alarm threshold** 200 或：[HUAWEI-GigabitEthernet1/0/0] **arp-miss anti-attack rate-limit alarm threshold** 100 或：[HUAWEI-Vlanif100]**arp-miss anti-attack rate-limit alarm threshold** 100	（可选）在全局、VLAN 或接口下配置 ARP Miss 消息限速丢弃告警阈值，取值范围为 1～16 384 的整数。通过本命令可以配置告警阈值，当设备忽略因超过 ARP Miss 消息限速值的 ARP Miss 消息数超过告警阈值时，设备将以告警的方式通知网络管理员 缺省情况下，ARP Miss 消息限速丢弃告警阈值为 100，可用 **undo arp-miss anti-attack rate-limit alarm threshold** 命令恢复 ARP Miss 消息限速丢弃告警阈值为缺省值

16.2.6 配置临时ARP表项的老化时间

为了控制设备根据 ARP Miss消息生成大量临时的 ARP表项，可以通过配置临时ARP表项的老化时间来控制ARP Miss消息的触发频率。这样，在送ARP Misss消息时，在老化时间超时前如果设备没有收到对应的ARP应答报文，则匹配对应临时ARP表项的 IP报文将被丢弃，并且不会再触发新的 ARP Miss消息；而在设备收到对应的 ARP应答报文后，会生成正确的ARP表项来替换临时ARP表项。

但仅配置老化时间还不够，因为在对应临时ARP表项的老化时间超时后，设备虽然会清除该临时ARP表项，但此时如果设备转发IP报文再次匹配不到对应的ARP表项时，则又会重新触发ARP Miss消息并生成临时ARP表项，如此循环重复。这时可调大对应IP报文的临时ARP表项的老化时间，以减小设备的ARP Miss消息的触发频率，从而减小攻击对设备的影响。

临时 ARP 表项的老化时间是在具体的 VLANIF 接口视图下配置的，以配置对应VLAN接口所在IP网段的临时ARP表项的老化时间，具体配置步骤如表16-8所示。

说明

本项 ARP 安全特性，在 S2700 系列中的 S2700SI 子系列和除 S2700-52P-EI、 S2700-52P-PWR-EI以外的S2700EI子系列不支持；其他系列均支持。

表16-8 临时ARP表项的老化时间的配置步骤

步骤	命令	说明
1	**system-view** 例如：<HUAWEI> **system-view**	进入系统视图
2	**interface vlanif** *interface-number* 例如：[HUAWEI] **interface vlanif** 10	键入要配置临时 ARP 表项的老化时间的 VLANIF 接口，进入 VLANIF 接口视图
3	**arp-fake expire-time** *expire-time* 例如：[HUAWEI-Vlanif10]**arp-fake expire-time** 100	配置临时 ARP 表项的老化时间，取值范围为 1～36 000 的整数秒 缺省情况下，临时 ARP 表项的老化时间是 1s，可用 **undo arp-fake expire-time** 命令恢复对应 VLAN 所在 IP 网段的临时 ARP 表项的老化时间为缺省值

16.2.7 配置ARP表项严格学习

如果大量用户在同一时间段内向设备发送大量ARP报文，或者攻击者伪造正常用户的ARP报文发送给设备，则会造成如下危害。

（1）设备因处理大量ARP报文而导致CPU负荷过重，同时设备学习大量的ARP报文可能导致设备ARP表项资源被无效的ARP表项耗尽，造成合法用户的ARP报文不能继续生成ARP表项，导致用户无法正常通信。

（2）伪造的ARP报文将错误地更新设备ARP表项，导致合法用户无法正常通信。

为避免上述危害，可以在网关设备上配置ARP表项严格学习功能。配置该功能后，只有本设备主动发送的ARP请求报文的应答报文才能触发本设备学习ARP，其他设备主动向本设备发送的ARP报文不能触发本设备学习ARP，这样，可以拒绝大部分的ARP报文攻击。

ARP 表项严格学习功能可在全局和 VLANIF 接口视图下进行配置，具体如表 16-9所示。如果全局使能该功能，则设备的所有接口均进行ARP表项严格学习；如果VLANIF接口下使能该功能，则只有该接口进行ARP 表项严格学习。当同时在全局和 VLANIF接口视图下进行配置时，VLANIF接口下配置的优先级高于全局配置的优先级。

说明

本项 ARP 安全特性，在 S2700 系列中的 S2700SI 子系列和除 S2700-52P-EI、 S2700-52P-PWR-EI以外的S2700EI子系列不支持；其他系列均支持。

表16-9 ARP表项严格学习的配置步骤

步骤	命令	说明
1	**system-view** 例如：<HUAWEI> **system-view**	进入系统视图
2	**arp learning strict** 例如：[HUAWEI] **arp learning strict**	配置全局 ARP 表项严格学习功能，使设备只学习自己发送的 ARP 请求报文的应答报文 缺省情况下，S2700/3700/S5700SI 和 S5700EI 已使能，其他系列均没有使能，可用 **undo arp learning strict** 命令将 ARP 表项严格学习功能恢复为缺省值

（续表）

步骤	命令	说明
3	**interface vlanif** *interface-number* 例如：[HUAWEI] **interface vlanif** 10	（可选）键入要配置 ARP 表项严格学习功能的 VLANIF 接口，进入 VLANIF 接口视图
4	**arp learning strict** { **force-enable** | **force-disable** | **trust** } 例如：[HUAWEI-Vlanif10] **arp learning strict trust**	（可选）配置 VLANIF 接口的 ARP 表项严格学习功能。命令中的选项说明如下。 ● **force-enable**：多选一选项，使能 ARP 严格学习功能 ● **force-disable**：多选一选项，去使能 ARP 严格学习功能 ● **trust**：多选一选项，ARP 严格学习功能与全局配置保持一致 缺省情况下，S2700/3700/S5700SI 和 S5700EI 已使能，其他系列均没有使能，可用 **undo arp learning strict** 命令将对应 VLANIF 接口的 ARP 严格学习功能配置与全局配置保持一致

说明

由于有些用户主机上安装的防火墙会阻止其收到ARP请求时发送ARP应答，所以使能ARP表项严格学习功能后，如果设备上触发了ARP Miss消息，则设备主动发出的ARP请求将无法得到该用户的ARP应答，从而使设备无法学习到该用户的ARP。在这种场景下，如果仅是个别用户出现该问题，则可以为其配置静态ARP；如果该问题在用户中非常普遍，则建议去使能ARP表项严格学习功能。

16.2.8 配置基于接口的ARP表项限制

为了防止当一个接口（可以是二层物理接口、Eth-Trunk接口、VLANIF接口、三层物理子接口和端口组）所接入的某一用户主机发起ARP攻击时导致整个设备的ARP表资源被耗尽，可以在指定接口下配置接口能够学习到的最大动态ARP表项数目。当指定接口下的动态ARP表项达到允许学习的最大数目后，将不允许新增动态ARP表项。具体的配置步骤如表16-10所示。

说明

本项 ARP 安全特性，在 S2700 系列中的 S2700SI 子系列和除 S2700-52P-EI、S2700-52P-PWR-EI系列以外的S2700EI子系列产品不支持；S2710SI、S2700-52P-EI和S2700-52P-PWR-EI子系列仅支持本项特性 （但不支持在物理子接口下配置）；其他S系列均支持本项特性（但在S5700系列中仅S5710EI和S5700HI子系列支持在子接口下配置，S9300/9300E系列不支持在子接口下配置）。

表16-10 基于接口的ARP表项限制的配置步骤

配置任务	步骤	命令	说明
公共配置	1	**system-view** 例如：<HUAWEI> **system-view**	进入系统视图
在物理接口或 Eth-Trunk 接口视图下配置	2	**interface** *interface-type interface-number* 例如：[HUAWEI] **interface gigabitethernet 1/0/0**	（可选）键入要配置 ARP 表项限制功能的物理接口或 Eth-Trunk 接口或端口组，进入接口或者端口组视图

（续表）

配置任务	步骤	命令	说明
在物理接口或 Eth-Trunk 接口视图下配置	3	**arp-limit vlan** *vlan-id1* [**to** *vlan-id2*] **maximum** *maximum* 例如：[HUAWEI-GigabitEthernet1/0/0] **arp-limit vlan** 10 **maximum** 20	（可选）配置基于物理主接口、Eth-Trunk 接口或端口组的 ARP 表项限制。命令中的参数说明如下。 （1）*vlan-id1* [**to** *vlan-id2*]：指定限制 ARP 学习的 VLAN，限制接口从该 VLAN 内能够学习到的最大动态 ARP 表项数目，取值范围均为 1～4 094，但 *vlan-id2* 必须大于或等于 *vlan-id1*。该参数必须且只能在二层接口下使用 （2）*maximum*：接口能够学习到的最大动态 ARP 表项数目，S2700/3700 系列的取值范围是 1～1 024 的整数，S5700SI 系列的取值范围为 1～2 048 的整数，S5700EI 系列取值范围为 1～8 192 的整数，S5700HI/7700/9300/9300E/9600 系列的取值范围为 1～16 384 的整数，S5710EI 系列的取值范围为 1～16 384 的整数，S5700LI/5700S-LI 系列的取值范围为 1～256 的整数，S6700 系列的取值范围为 1～8 192 的整数 缺省情况下，在规格范围内，设备对接口能够学习到的最大动态 ARP 表项数目没有限制，可用 **undo arp-limit vlan** *vlan-id1* [**to** *vlan-id2*] 命令删除对应接口下指定 ARP 表项限制配置
在 VLNIF 接口视图下配置	4	**quit** 例如：[HUAWEI-GigabitEthernet1/0/0] **quit**	（可选）退出物理或者 Eth-Trunk 接口视图，返回系统视图
	5	**interface vlanif** *interface-number* 例如：[HUAWEI]**interface vlanif 10**	（可选）键入要配置 ARP 表项限制功能的 VLAN 接口，进入 VLANIF 接口视图
	6	**arp-limit maximum** *maximum* 例如：[HUAWEI-Vlanif10] **rp-limit maximum** 20	（可选）配置基于 VLANIF 接口的 ARP 表项限制。参数 *maximum* 的取值范围参见前面第 3 步中该参数说明 缺省情况下，在规格范围内，设备对 VLANIF 接口能够学习到的最大动态 ARP 表项数目没有限制，可用 **undo arp-limit** 命令删除对应 VLANIF 接口下的指定 ARP 表项限制配置
在物理子接口视图下配置	7	**quit** 例如：[HUAWEI-Vlanif10] **quit**	（可选）退出 VLANIF 接口视图，返回系统视图
	8	**interface** *interface-type interface-number* [.*subnumber*] 例如：[HUAWEI] **interface** gigabitethernet 1/0/1.1	（可选）键入要配置 ARP 表项限制功能的物理子接口，进入子接口视图
	9	**arp-limit vlan** *vlan-id1* [**to** *vlan-id2*] **maximum** *maximum* 例如：[HUAWEI-GigabitEthernet1/0/1.1] **arp-limit vlan** 10 **maximum** 20	（可选）配置基于物理子接口的 ARP 表项限制。命令中的参数参见前面第 3 中对应参数说明 缺省情况下，在规格范围内，设备对子接口能够学习到的最大动态 ARP 表项数目没有限制，可用 **undo arp-limit** 命令删除对应 VLANIF 接口下的指定 ARP 表项限制配置

16.2.9 配置免费ARP报文主动丢弃

由于发送免费ARP报文的用户主机并不需要经过身份验证，任何一个用户主机都可以发送免费ARP报文，这样就引入了两个问题。

（1）如果网络中出现大量的免费 ARP 报文，设备会因为处理这些报文而导致 CPU负荷过重，从而不能正常处理合法的ARP报文。

（2）如果设备处理的免费 ARP 报文是攻击者伪造的，会造成设备错误地更新 ARP表项，导致合法用户的通信流量发生中断。

为了解决以上问题，在确认攻击来自免费ARP报文之后，可以在网关设备上使能免费ARP报文主动丢弃功能，使网关设备直接丢弃所收到的免费ARP报文。

丢弃免费 ARP 报文功能可以在全局和 VLANIF 接口下使能，具体配置步骤如表16-11所示。全局使能该功能，则设备的所有接口都丢弃收到的免费ARP报文。VLANIF接口下使能该功能，则只有该接口丢弃收

到的免费ARP报文。一般在用户侧的VLANIF接口下配置免费ARP报文主动丢弃功能。

说明

本项ARP安全特性仅S7700/9300/9300E/9700系列支持。

表16-11 免费ARP报文主动丢弃的配置步骤

步骤	命令	说明
1	**system-view** 例如：<HUAWEI> **system-view**	进入系统视图
2	**interface vlanif** *interface-number* 例如：[HUAWEI] **interface vlanif** 10	键入要配置临时 ARP 表项的老化时间的 VLANIF 接口，进入 VLANIF 接口视图。**在系统视图下使能免费 ARP 报文主动丢弃功能无需执行此步骤**
3	**arp anti-attack gratuitous-arp drop** 例如：[HUAWEI]**arp anti-attack gratuitous-arp drop** 或：[HUAWEI-Vlanif10] **arp anti-attack gratuitous-arp drop**	使能免费 ARP 报文主动丢弃功能 缺省情况下，没有使能免费ARP报文主动丢弃功能，可用 **undo arp anti-attack gratuitous-arp drop** 命令去使能丢弃免费 ARP 报文功能

16.3 配置防ARP欺骗攻击

ARP表项攻击、网关攻击、中间人攻击是ARP欺骗攻击的主要应用场景。表16-12列出了针对不同ARP欺骗攻击类型所提供的不同解决方案，以增强网络抗击ARP欺骗攻击的能力。在这些配置防ARP欺骗攻击安全特性中，各项配置均是并列关系，无严格配置顺序，也并不是要求配置所有的ARP安全特性方案，用户可根据需要选择配置一种或者多种方案。

表16-12 防ARP欺骗攻击的类型及解决方案

攻击类型	攻击方式	解决方案
ARP 表项攻击	ARP 欺骗攻击一般都是通过修改 ARP 表项来完成的。	增强 ARP 表项的自我保护功能，可采用以下具体的解决方案。 （1）ARP 表项固化：使设备在第一次学习到 ARP 之后，不再允许用户更新此 ARP 表项或只能更新此 ARP 表项的部分信息，或者通过发送 ARP 请求报文的方式进行确认，以防止攻击者伪造ARP报文修改正常用户的ARP表项内容。**具体见 16.3.1 节** （2）ARP 报文合法性检查：使设备对收到的 ARP 报文进行以太网数据帧首部中的源MAC地址和ARP报文数据区中的源 MAC 地址的一致性检查，如果两者不一致，则直接丢弃该 ARP 报文，以免非法修改或创建 ARP 表项，否则允许该 ARP 报文通过。**具体见 16.3.6 节** （3）ARP 表项严格学习：只学习自己发送的 ARP 请求报文的应答报文. **具体参见 16.2.7 节** （4）ARP 报文内 MAC 地址一致性检查：使网关设备在进行 ARP 学习前对 ARP 报文进行检查。如果以太网数据帧首部中的源/目的 MAC 地址和 ARP 报文数据区中的源/目的 MAC 地址不同，则认为是攻击报文，将其丢弃；否则，继续进行 ARP 学习。**具体见 16.3.5 节** （5）DHCP 触发 ARP 学习：当 DHCP 服务器给用户分配 IP 地址时，使设备回应用户 DHCP ACK 报文成功后，会取用户的 MAC 地址，生成该 IP 地址对应的 ARP 表项。这样可以省掉设备学习用户主机 ARP 的过程，避免攻击者通过 ARP 报文对 ARP 表项的攻击. **具体见 16.3.7 节**
网关攻击	攻击者仿冒网关地址，发送ARP报文头的源IP地址是网关地址的 ARP 报文，从而使主机修改网关的 MAC 地址为攻击者的 MAC 地址，需要发送给原来网关的报文就发送给攻击者了	可采用以下具体的解决方案： （1）ARP 防网关冲突：使配置作为网关的设备丢弃 ARP 报文头的源 IP 地址是自己 IP 地址的 ARP 报文。但是这种防攻击策略只适用于所有主机的 ARP 报文必须通过网关转发的场景。**具体见 16.3.3 节** （2）发送免费 ARP 报文：用来定期更新合法用户的 ARP 表项，使得合法用户 ARP 表项中记录的是正确的网关地址映射关系。**具体见 16.3.4 节**
中间人 攻击	中间人攻击会同时修改主机和网关的信息。 （1）修改主机上的网关信息：攻击者仿冒网关地址，发送ARP报文头的源IP地址是网关地址的 ARP 报文，使主机修改网关的 MAC 地址为攻击者的 MAC 地址 （2）修改网关上的主机信息：攻击者仿冒主机地址，发送ARP报文头的源IP地址是主机地址的 ARP 报文，使网关修改主机的 MAC 地址为攻击者的 MAC 地址	可采用"动态 ARP 检测"解决方案：当设备收到 ARP 报文时，将此 ARP 报文的源 IP、源 MAC、收到 ARP 报文的接口及 VLAN 信息和 DHCP Snooping 绑定表的信息进行比较，如果信息匹配，则认为是合法用户，允许此用户的 ARP 报文通过，否则认为是攻击，丢弃该 ARP 报文。本功能仅适用于 DHCP Snooping 场景，适用于所有主机的 ARP 报文必须通过网关转发的场景。**具体见 16.3.2 节**

16.3.1 配置ARP表项固化

为了防止ARP地址欺骗攻击，可以配置ARP表项固化功能，使欺骗类ARP报文不能修改原来ARP表项。以下3种ARP表项固化模式适用于不同的应用场景，且是互斥关系。

（1）**fixed-mac** 方式：这种固化模式是以报文中源**MAC**地址与**ARP**表中现有对应**IP**地址的表项中的**MAC**地址是否匹配为审查的关键依据 。当这两个MAC地址不匹配时，则直接丢弃该 ARP 报文；如果这两个 MAC 地址是匹配的，但是报文中的接口或VLAN信息与ARP表中对应表项不匹配时，则可以更新对应ARP表项中的接口和VLAN信息。这种模式适用于静态配置IP地址，但网络存在冗余链路（这样可以改变出接口和**VLAN** ）的情况。当链路切换时，ARP表项中的接口信息可以快速改变。

（2）**fixed-all** 方式：这种固化模式是仅当ARP报文对应的MAC地址、接口、VLAN信息和ARP表中对应表项的信息完全匹配时，设备才可以更新ARP表项的其他内容。这种模式匹配最严格 ，适用于静态配置IP地址，网络没有冗余链路（这样不可以改变出接口和**VLAN** ），且同一IP地址用户不会从不同接口接入

的情况。

（3）**send-ack** 方式：这种模式是当设备收到一个涉及 MAC 地址、VLAN、接口修改的ARP报文时，不会立即更新ARP表项，而是先向待更新的ARP表项现有MAC地址对应的用户发送一个单播的ARP请求报文，再根据用户的确认结果决定是否更新ARP表项中的MAC地址、VLAN和接口信息。此方式适用于动态分配IP地址，有冗余链路的网络。

可在全局和VLANIF接口下配置ARP表项固化功能，具体配置步骤如表16-13所示。全局配置该功能后，缺省设备上所有接口的 ARP 表项固化功能均已使能。当全局和VLANIF接口下同时配置了该功能时，VLANIF接口下的配置优先生效。

说明

本项 ARP 安全特性，在 S2700 系列中的 S2700SI 子系列和除 S2700-52P-EI、 S2700-52P-PWR-EI以外的S2700EI子系列不支持；其他系列均支持。

表16-13 ARP表项固化的配置步骤

步骤	命令	说明
1	**system-view** 例如：<HUAWEI> **system-view**	进入系统视图
2	**interface vlanif** *interface-number* 例如：[HUAWEI] **interface vlanif** 10	（可选）键入要配置 ARP 表项严格学习功能的 VLANIF 接口，进入 VLANIF 接口视图。**在系统视图下配置 ARP 表项固化功能无需执行此步骤**
3	**arp anti-attack entry-check** { **fixed-mac** | **fixed-all** | **send-ack** } **enable** 例如：[HUAWEI]**arp anti-attack entry-check fixed-mac** 或[HUAWEI-Vlanif10]**arp anti-attack entry-check send-ack**	在全局或 VLANIF 接口下配置 ARP 表项固化功能。命令中的选项说明如下。 （1）**fixed-mac**：多选一选项，指定按固定 MAC 模式运行 ARP 防欺骗功能。固定 MAC 指的是不允许通过 ARP 学习对 MAC 地址进行修改，但允许对 VLAN 和接口信息进行修改 （2）**fixed-all**：多选一选项，指定按固定所有参数的模式运行 ARP 防欺骗功能。固定所有参数指的是对动态 ARP 和已解析的静态 ARP，MAC、VLAN 和接口信息均不允许修改 （3）**send-ack**：多选一选项，指定按查询确认模式运行 ARP 防欺骗功能。查询确认指的是设备收到一个涉及 MAC 地址、VLAN、接口修改的 ARP 报文时，不会立即进行修改，而是先记录发送请求的表项信息，对原 ARP 表中与此 ARP 报文中的 MAC 地址对应的用户发一个单播确认，在收到 ACK 后删除该表项 缺省情况下，没有使能 ARP 防地址欺骗功能，可用 **undo arp anti-attack entry-check** [**fixed-mac** | **fixed-all** | **send-ack**] **enable** 命令去使能对应的 ARP 防地址欺骗功能

（续表）

16.3.2 配置动态ARP检测

为了防御中间人攻击，避免合法用户的数据被中间人窃取，可以使能动态ARP检测功能，仅适用于启用了**DHCP Snooping**的场景 。这样，设备会将ARP报文中的源 IP、源MAC、接口、VLAN信息和DHCP Snooping中的绑定表（或者手动添加静态绑定表）信息进行比较，如果匹配，说明发送该ARP报文的用户是合法用户，允许此用户的ARP报文通过，否则就认为是攻击，丢弃该ARP报文。

说明

设备使能DHCP Snooping功能后，当DHCP用户上线时设备会自动生成DHCP Snooping绑定表；对于静态配置 IP地址的用户，设备不会生成DHCP Snooping绑定表，所以需要手动添加静态绑定表。可用 user-bind static{ ip-address start-ip [to end-ip] &<1-10> |mac-addressmac-address } * [interface interface-type interface-number] [vlanvlan-id [ce-vlan ce-vlan-id]] 命令配置 IP地址、MAC地址、接口和内/外层VLAN的静态用户绑定表项。

如果希望仅匹配绑定表中某一项或某两项内容的特殊ARP报文也能够通过，则可以配置对ARP报文进行绑定表匹配检查时只检查某一项或某两项内容。如果希望设备在丢弃的不匹配绑定表的ARP报文数量较多时能够以告警的方式提醒网络管理员，则还可以使能动态ARP检测丢弃报文告警功能。这样，当丢弃的ARP报文数超过告警阈值时，设备将产生告警。

可在接口（包括物理接口、**Eth-Trunk**接口、**VLANIF**接口和端口组 ）视图或VLAN视图下配置动态ARP检测功能，具体配置步骤如表16-14所示。在接口视图下使能时，则对该接口收到的所有ARP报文进行绑定表匹配检查；在VLAN视图下使能时，则对加入该VLAN的接口收到的属于该VLAN的ARP报文进行绑定表匹配检查。

说明

本项ARP安全特性，S2700SI和S5700SI子系列交换机不支持，其他S系列均支 持。除S2700-52P-EI、S2700-52P-PWR-EI以外的S2700EI子系列产品仅支持本项特性 。

另外，当同时在VLAN和加入该VLAN的接口下配置了动态ARP检测功能时，设备会先按照接口下配置的检查选项对ARP报文进行绑定表匹配检查，如果ARP报文检查通过，设备再根据VLAN下配置的检查选项进行检查。

由于动态ARP检测丢弃报文告警功能针对的是接口下DAI功能丢弃的ARP报文数告警统计，因此建议不要同时在VLAN视图下配置命令 arp anti-attack check user-bind enable 以及在加入该 VLAN 的接口视图下配置命令 arp anti-attack check user-bind alarm enable，避免VLAN下的DAI功能可能造成实际丢包数目和接口DAI告警统计值之间的偏差。

表16-14 动态ARP检测的配置步骤

步骤	命令	说明
1	**system-view** 例如：<HUAWEI> **system-view**	进入系统视图
2	**interface** *interface-type interface-number* 例如：[HUAWEI] **interface gigabitethernet 1/0/0** 或 **vlan** *vlan-id* 例如：[HUAWEI] **vlan** 10	键入要配置动态 ARP 检测功能的物理接口、Eth-Trunk 接口、VLNIF 接口或端口组，进入接口或者端口组视图，或键入要配置动态 ARP 检测功能的 VLAN，进入 VLAN 视图
3	**arp anti-attack check user-bind enable** 例如：[HUAWEI-GigabitEthernet1/0/0] **arp anti-attack check user-bind enable** 或[HUAWEI-Vlan10]**arp anti-attack check user-bind enable**	在以上接口或者 VLAN 下使能动态 ARP 检测功能（即对 ARP 报文进行绑定表匹配检查功能）。缺省情况下，没有使能动态 ARP 检测功能。
4	**arp anti-attack check user-bind check-item { ip-address | mac-address | vlan }** 例如：[HUAWEI-GigabitEthernet1/0/0] **arp anti-attack check user-bind check-item ip-address**	（可选）在以上接口下配置对 ARP 报文进行绑定表匹配检查的检查项。命令中的选项说明如下。 （1）**ip-address**：可多选项，指定 ARP 报文绑定表匹配检查时检查 IP 地址 （2）**mac-address**：可多选项，指定 ARP 报文绑定表匹配检查时检查 MAC 地址 （3）**vlan**：可多选项，指定 ARP 报文绑定表匹配检查时检查 VLAN 信息 缺省情况下，对 ARP 报文的 IP 地址、MAC 地址和 VLAN 信息都进行检查，可用 **undo arp anti-attack check user-bind check-item** 命令恢复 ARP 报文绑定表匹配检查项为缺省值
5	**arp anti-attack check user-bind check-item { ip-address | mac-address | interface }*** 例如：[HUAWEI-vlan10] **arp anti-attack check user-bind check-item ip-address**	（可选）在以上 VLAN 下配置对 ARP 报文进行绑定表匹配检查的检查项。命令中的选项说明如下。 （1）**ip-address**：可多选项，指定 ARP 报文绑定表匹配检查时检查 IP 地址 （2）**mac-address**：可多选项，指定 ARP 报文绑定表匹配检查时检查 MAC 地址 （3）**interface**：可多选项，指定 ARP 报文绑定表匹配检查时检查接口信息 缺省情况下，对 ARP 报文的 IP 地址、MAC 地址和接口信息都进行检查，可用 **undo arp anti-attack check user-bind check-item** 命令恢复 ARP 报文绑定表匹配检查项为缺省值

（续表）

步骤	命令	说明	
6	**arp anti-attack check user-bind alarm enable** 例如：[HUAWEI-GigabitEthernet1/0/0] **arp anti-attack check user-bind alarm enable**	（可选）在以上接口上使能动态 ARP 检测丢弃报文告警功能 缺省情况下，没有使能动态 ARP 检测丢弃报文告警功能，可用 **undo arp anti-attack check user-bind alarm enable** 命令去使能动态 ARP 检测丢弃报文告警功能	除 S2700-52P-EI 和 S2700-52P-PWR-EI 以外的 S2700EI 系列，以及 S5700LI 和 S5700S-LI 系列均不支持这两项配置，其他 S 系列支持这两项配置
7	**arp anti-attack check user-bind alarm threshold** *threshold* 例如：[HUAWEI-GigabitEthernet1/0/0] **arp anti-attack check user-bind alarm threshold** 200	（可选）在以上接口上配置动态 ARP 检测丢弃报文告警阈值，取值范围为 1～1 000 缺省情况下，动态 ARP 检测丢弃报文告警阈值为系统视图下 **arp anti-attack check user-bind alarm threshold** *threshold* 命令配置的值。如果系统视图下没有配置该值，则接口下缺省的告警阈值为 100	

16.3.3 配置ARP防网关冲突

如果攻击者仿冒网关，在局域网内部发送源IP地址是网关IP地址的ARP报文，就会导致局域网内其他用户主机的ARP表记录错误的网关地址映射关系。这样其他用户主机就会把发往网关的流量均发送给了攻击者，攻击者可轻易窃听到他们发送的数据内容，并且最终会造成这些用户主机无法访问网络。这就是我们最常见的一种ARP攻击类型——网关欺骗ARP攻击。

为了防范攻击者仿冒网关，当用户主机直接接入网关时（这是先决条件，其他情形不适用），可在网

关设备上使能ARP防网关冲突攻击功能。这样，当网关设备收到的ARP报文存在下列情况之一时，设备就会认为该ARP报文是与网关地址冲突的ARP报文，生成 ARP 防攻击表项，并在后续一段时间内丢弃该接口收到的相同 VLAN 和相同源MAC地址的ARP报文，以防止与网关地址冲突的ARP报文在VLAN内广播。

（1）ARP报文的源IP地址与报文入接口对应的VLANIF接口的IP地址（就是网关IP地址）相同。

（2）ARP报文的源IP地址是入接口的VRRP虚拟IP地址，但ARP报文源MAC地址不是VRRP虚拟MAC地址。

本项ARP安全特性，在S2700系列中，S2700SI和除S2700-52P-EI、S2700- 52P-PWR-EI以外的S2700EI系列，以及S5700LI和S5700S-LI系列不支持；其他S系列均支持。

说明

配置 ARP 防网关冲突的方法很简单，仅需在系统视图下执行 arp anti-attack gateway-duplicate enable命令即可。缺省情况下，没有使能ARP防网关冲突攻击功能，可用undo arp anti-attack gateway-duplicate enable命令去使能ARP防网关冲突攻击功能。

【经验之谈 】这种方法不是很适用，因为它只适用于用户主机直接与网关连接的情形，仅在网关上进行了ARP报文检查。在其他情形下，其他非网关设备照样不会对这类报文进行检查。建议配合下节将要介绍的在网关上定期发送ARP免费报文的方法。

16.3.4 配置发送ARP免费报文

如果有攻击者向其他用户发送仿冒网关的ARP报文，会导致其他用户的ARP表中记录错误的网关地址映射关系，从而造成其他用户的正常数据不能被网关接收。此时可以在网关设备上配置发送免费ARP报文的功能，用来定期更新合法用户的ARP表项，使得合法用户ARP表项中记录的是正确的网关地址映射关系。

可在网关设备上全局或VLANIF接口下配置发送免费ARP报文功能，具体配置步骤如表16-15所示。全局配置该功能后，则缺省设备上所有接口的发送ARP免费报文功能均已使能。当全局和VLANIF接口下同时配置了该功能时，VLANIF接口下的配置优先生效。

说明

本项ARP安全特性，S5700LI和S5700S-LI子系列不支持，其他S系列均支持。另外，本方案的有效性也不是很高，毕竟网关不可能总发送免费ARP报文，攻击者 还是有空子可钻的。建议与其他ARP特性方案配合使用。

表16-15 发送ARP免费报文的配置步骤

步骤	命令	说明
1	**system-view** 例如：<HUAWEI> **system-view**	进入系统视图
2	**interface vlanif** *interface-number* 例如：[HUAWEI] **interface vlanif** 10	（可选）键入要配置主动发送免费 ARP 报文的 VLANIF 接口，进入 VLANIF 接口视图。**在系统视图下配置主动发送免费 ARP 报文功能无需执行此步骤**
3	**arp gratuitous-arp send enable** 例如：[HUAWEI]**arp gratuitous-arp send enable** 或[HUAWEI-Vlanif10]**arp gratuitous-arp send enable**	使能主动发送免费 ARP 报文的功能 缺省情况下，主动发送免费 ARP 报文功能处于没有使能状态，可用 **undo arp gratuitous-arp send enable** 命令去使能主动发送免费 ARP 报文的功能
4	**arp gratuitous-arp send interval** *interval-time* 例如：[HUAWEI] **arp gratuitous-arp send interval** 1000 或[HUAWEI-Vlanif10] **arp g ratuitous-arp send interval 100**	配置主动发送免费 ARP 报文的时间间隔，取值范围为 1～86 400 的整数秒 缺省情况下，S2700/3700/5700/6700 系列发送免费 ARP 报文的时间间隔为 60s，S7700/9300/9300E/9700 系列发送免费 ARP 报文的时间间隔为 30s，可用 **undo arp gratuitous-arp send interval** 命令将发送免费 ARP 报文的时间间隔恢复为缺省值

16.3.5 配置ARP报文内MAC地址一致性检查

ARP报文内MAC地址一致性检查功能主要应用于网关设备上，可以防御以太网数据帧首部中的源/目的MAC地址和ARP报文数据区中的源/目的MAC地址不同的ARP攻击。配置该功能后，网关设备在进行ARP表项学习前将对ARP报文进行检查。如果以太网数据帧首部中的源/目的MAC地址和ARP报文数据区中的源/目的MAC地址不同，则认为是攻击报文，将其丢弃；否则，继续进行ARP学习。

说明

本项ARP安全特性，S2700/3700系列交换机不支持，且其他系列也只能在物理接口或者Eth-Trunk接口下配置，不支持在VLANIF接口和物理子接口上配置。当VLANIF接口收到ARP报文时，ARP报文内MAC地址一致性检查遵循成员接口下的检查规则；当物理 子接口收到ARP报文时，ARP报文内MAC地址一致性检查遵循主接口下的检查规则。

ARP报文内MAC地址一致性检查的配置步骤如表16-16所示。

表16-16 ARP报文内MAC地址一致性检查的配置步骤

步骤	命令	说明
1	**system-view** 例如：<HUAWEI> **system-view**	进入系统视图
2	**interface** *interface-type interface-number* 例如：[HUAWEI] **interface** gigabitethernet 1/0/1	（可选）键入要配置 ARP 报文内 MAC 地址一致性检查功能的**物理接口或者 Eth-Trunk 接口**，进入接口视图
3	**arp validate** { **source-mac** | **destination-mac** }* 例如：[HUAWEI-GigabitEthernet1/0/1] **arp validate source-mac destination-mac**	使能 ARP 报文内 MAC 地址一致性检查功能，即对以太网数据帧首部中的源/目的 MAC 地址和 ARP 报文数据区中的源/目的 MAC 地址进行一致性检查的功能。命令中的选项说明如下。 （1）**source-mac**：可多选选项，指定接口收到 ARP 报文时对以太网数据帧首部中的源 MAC 地址和 ARP 报文数据区中的源 MAC 地址进行一致性检查。此时，当接口收到 ARP 请求或者应答报文时，均只对报文中的源 MAC 地址进行一致性检查 （2）**destination-mac**：可多选选项，指定接口收到 ARP 报文时对以太网数据帧首部中的目的 MAC 地址和 ARP 报文数据区中的目的 MAC 地址进行一致性检查。此时，当接口收到 ARP 请求报文时，不对报文进行一致性检查，因为 ARP 请求报文是广播报文，目的地址是广播 MAC 地址；当接口收到 ARP 应答报文时，对报文中的目的 MAC 地址进行一致性检查 如果同时选择以上两可选项时，当接口收到 ARP 请求报文时，只对报文中的源 MAC 地址进行一致性检查；当接口收到 ARP 应答报文时，对报文中的源/目的 MAC 地址都进行一致性检查 缺省情况下，不对以太网数据帧首部中的源/目的 MAC 地址和 ARP 报文数据区中的源/目的 MAC 地址进行一致性检查，可用 **undo arp validate** { **source-mac** | **destination-mac** }*命令去使能对应的 ARP 报文内 MAC 地址一致性检查功能

16.3.6 配置ARP报文合法性检查

为了防止非法ARP报文的攻击，可以在接入设备或网关设备 上 配置ARP报文合法性检查功能，用来对MAC地址和IP地址不合法的ARP报文进行过滤。设备提供以下三种可以任意组合的检查项配置。

（1）IP地址检查：设备会检查ARP报文中的源IP和目的IP地址，全0、全1或者组播IP地址都是不合法的，需要丢弃。对于ARP应答报文，源IP和目的IP地址都进行检查；对于ARP请求报文，只检查源IP地址。

（2）源MAC地址检查：设备会检查ARP报文中的源MAC地址和以太网数据帧首部中的源MAC地址是否一致，一致则认为合法，否则丢弃报文。

（3）目的MAC地址检查：设备会检查ARP应答报文中的目的MAC地址是否和以太网数据帧首部中的目的MAC地址一致，一致则认为合法，否则丢弃报文。

说明

其实后面两项MAC地址的检查与上节介绍的ARP报文内MAC地址一致性检查 方法是一样的。但是，通常ARP报文中源MAC地址和以太网数据帧首部中的源MAC 地址不一致的ARP报文，以及目的MAC地址和以太网数据帧首部中的目的MAC地址不一致的ARP应答报文均是ARP协议允许的ARP报文。因此，只有在网络管理员发现攻击产生后，通过报文头获取方式定位、确定了是由于对应项不一致的ARP报文导致的攻击，才能指定ARP报文合法性，检查时需要检查源MAC地址和目的MAC地址。

本项ARP安全特性，除S2700-52P-EI、S2700-52P-PWR-EI以外的S2700EI子系列，以及S5700LI和S5700S-LI子系列不支持，其他S系列均支持。

ARP 报文合法性检查的配置方法很简单，仅需在系统视图下配置 arp anti-attack packet-check { ip |dst-mac | sender-mac }*命令，使能 ARP报文合法性检查功能，并指定ARP报文合法性检查项。命令中的选项说明如下。

（1）**ip**： 可多选选项，对应前面介绍的“IP地址检查”方式，指定在进行ARP报文合法性检查时检查IP地址。S2700/3700/9300/9300系列不支持该选项。

（2）**dst-mac**： 可多选选项，对应前面介绍的“目的MAC地址检查”方式，指定在进行ARP报文合法性检查时检查目的MAC地址。**S2700/3700/9300/9300**系列不支持该选项 。

（3）**sender-mac**：可多选选项，对应前面介绍的“源 MAC 地址检查”方式，指定在进行ARP报文合法性检查时检查源MAC地址。

缺省情况下，没有使能 ARP 报文合法性检查功能，可用 undo arp anti-attack packet-check [ip |dst-mac | sender-mac] *命令去使能ARP报文合法性检查功能。

16.3.7 配置DHCP触发ARP学习

在DHCP用户场景下，当DHCP用户数目很多时，设备进行大规模ARP表项的学习和老化会对设备性能和网络环境形成冲击。为了避免此问题，可以在网关设备上使能DHCP触发ARP学习功能。当DHCP服务器给用户分配了IP地址，网关设备会根据VLANIF接口上收到的DHCP ACK（确认）报文直接生成该用户的ARP表项。但DHCP触发ARP学习功能生效的前提是已在网关设备上通过dhcp snooping enable命令使能了DHCP Snooping功能。

说明

在VRRP和DHCP Relay组合场景下，VRRP主备设备上都不能再配置命令dhcp snooping enable和 arp learning dhcp-trigger。网关设备上还可同时部署动态ARP检测功能（参见本章16.3.2节），防止DHCP用户的ARP表项被伪造的ARP报文恶意修改。

本项ARP安全特性，在S2700系列中的S2700SI和S2710SI子系列，除S2700-52P-EI、S2700-52P-PWR-EI以外的S2700EI子系列，以及S5700LI和S5700S-LI子系列不支持；其他S系列均支持。

DHCP触发ARP学习的配置方法也很简单，就是在对应的VLANIF接口视图下执行arp learning dhcp-trigger命令使能DHCP触发ARP学习功能。缺省情况下，没有使能DHCP触发ARP学习功能，可用undoarp learning dhcp-trigger命令去使能DHCP触发ARP学习功能。

16.4 ARP安全配置管理

可使用以下display任意视图命令管理防ARP泛洪攻击配置。

（1）display arp anti-attack configuration {arp-rate-limit | arpmiss-rate-limit | arp- speed-limit | arpmiss-speed-limit |entry-check | gateway-duplicate | log-trap-timer |packe t-check | all }：查看ARP防攻击配置。

（2）display arp-limit [interface interface-type interface-number] [vlan vlan-id]：查看接口可以学习到的动态ARP表项数目的最大值。

（3）display arp learning strict：查看全局和所有VLANIF接口上的ARP表项严格学习情况。可使用以下display任意视图命令管理防ARP欺骗配置。

（1）display arp anti-attack configuration check user-bind interface interface-type interface-number：查看接口下ARP报文检查相关的配置。

（2）display arp anti-attack gateway-duplicate item：查看ARP防网关冲突攻击表项。

可使用以下display任意视图命令维护ARP安全包括监控ARP运行情况。

（1）display arp packet statistics：查看ARP处理的报文统计数据。

（2）display arp anti-attack statistics check user-bind interface interface-type interface-number：查看接口下进行ARP报文绑定表匹配检查的ARP报文丢弃计数。

（3）display arp anti-attack packet-check statistics：查看ARP报文合法性检查过程中被过滤的非法ARP报文统计数据。

（4）display arp anti-attack arpmiss-record-info [ip-address]：查看ARP Miss消息限速触发时的相关信息。

可使用以下reset用户视图命令清除ARP报文统计信息、清除ARP报文丢弃计数以及配置对潜在的ARP攻击行为发送日志和告警。清除统计信息后，以前的统计信息将无法恢复，务必仔细确认。

（1）reset arp packet statistics：清除ARP报文的统计信息。

（2）reset arp anti-attack statistics check user-bind interface interface-type interface-number：清除由于不匹配绑定表而丢弃的ARP报文计数。

（3）reset arp anti-attack statistics rate-limit：清除由于ARP报文超过速率限制阈值而被丢弃的计数。

16.5 配置示例

本节将以两个具体的配置示例介绍前面介绍的ARP安全配置。

16.5.1 ARP安全综合功能配置示例

本示例拓扑结构如图 16-1 所示，Switch 作为网关通过 GE1/0/3 接口连接一台服务器，通过GE1/0/1和GE1/0/2接口分别连接VLAN10和VLAN20下的用户。网络中存在以下ARP威胁，现希望能够防止这些ARP攻击行为，为用户提供更安全的网络环境和更稳定的网络服务。

（1）攻击者向Switch发送伪造的ARP报文和伪造的免费ARP报文进行ARP欺骗攻击，恶意修改Switch上的ARP表项，造成其他用户无法正常接收数据报文。

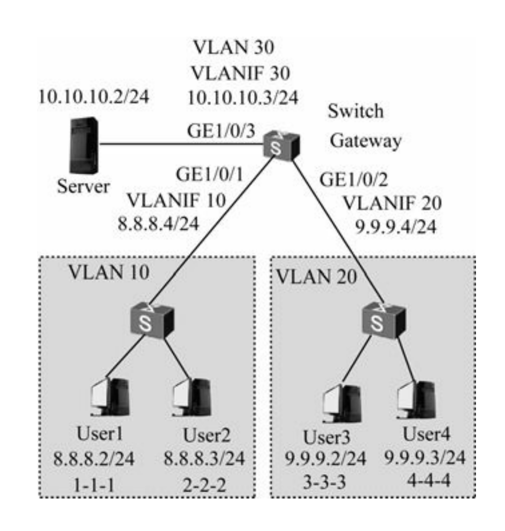


图16-1 ARP安全功能配置示例拓扑结构

（2）攻击者发出大量目的 IP 地址不可达的IP报文进行ARP泛洪攻击，造成Switch的CPU负荷过重。

（3）用户User1构造大量源IP地址变化、MAC地址固定的ARP报文进行ARP泛洪攻击，造成Switch的ARP表资源被耗尽以及CPU进程繁忙，影响正常业务的处理。

（4）用户User3构造大量源IP地址固定的ARP 报文进行 ARP 泛洪攻击，造成 Switch 的CPU进程繁忙，影响到正常业务的处理。

1．基本配置思路分析

针对这样的配置环境，首先要分析网络中存 在哪些ARP方面的安全隐患，然后有针对性地根据本章前面表16-2和表16-11所给出的可用解决方案选择对应的解决方案。针对示例中介绍的几种ARP威胁，对应可采用的解决方案如下。

（1）配置ARP表项严格学习功能和ARP表项固化功能，实现防止伪造的ARP报文错误地更新 Switch 的ARP 表项；配置免费 ARP 报文主动丢弃功能（仅适用于 **S7700/9300/9300E/9700**系列交换机 ），实现防止伪造的免费ARP报文错误地更新设备ARP表项。

（2）配置根据源 IP地址进行ARP Miss消息限速，实现防止用户侧存在攻击者发出大量目的 IP地址不可达的 IP报文触发大量ARP Miss消息，形成ARP泛洪攻击。同时需要保证Switch可以正常处理服务器发出的大量此类报文，避免因丢弃服务器发出的大量此类报文而造成网络无法正常通信。

（3）配置基于接口的ARP表项限制以及根据源MAC地址进行ARP限速，实现防止User1发送的大量源IP地址变化MAC地址固定的ARP报文形成的ARP泛洪攻击，避免Switch的ARP表资源被耗尽，并避免CPU进程繁忙。

（4）配置根据源IP地址进行ARP限速，实现防止User3发送的大量源IP地址固定的ARP报文形成的ARP泛洪攻击，避免Switch的CPU进程繁忙。

2．具体配置步骤

（1）批量创建VLAN10、VLAN20和VLAN30，并将GE1/0/1接口加入VLAN10中， GE1/0/2接口加入VLAN20中， GE1/0/3接口加入VLAN30中。

```
<HUAWEI>system-view
[HUAWEI] vlan batch 10 20 30
[HUAWEI] interface gigabitethernet 1/0/1
[HUAWEI-GigabitEthernet1/0/1] port link-type trunk
```

[HUAWEI-GigabitEthernet1/0/1] port trunk allow-pass vlan 10

[HUAWEI-GigabitEthernet1/0/1] quit

[HUAWEI] interfacegigabitethernet 1/0/2

[HUAWEI-GigabitEthernet1/0/2] port link-type trunk

[HUAWEI-GigabitEthernet1/0/2] port trunk allow-pass vlan 20

[HUAWEI-GigabitEthernet1/0/2] quit

[HUAWEI] interface gigabitethernet 1/0/3

[HUAWEI-GigabitEthernet1/0/3] port link-type trunk

[HUAWEI-GigabitEthernet1/0/3] port trunk allow-pass vlan 30

[HUAWEI-GigabitEthernet1/0/3] quit

（2）创建接口 VLANIF10、VLANIF20、VLANIF30，并按中标注配置各 VLANIF接口的IP地址。

[HUAWEI] interface vlanif 10

[HUAWEI-Vlanif10] ip address 8.8.8.4 24

[HUAWEI-Vlanif10] quit

[HUAWEI] interface vlanif 20

[HUAWEI-Vlanif20] ip address 9.9.9.4 24s

[HUAWEI-Vlanif20] quit

[HUAWEI] interface vlanif 30

[HUAWEI-Vlanif30] ip address 10.10.10.3 24

[HUAWEI-Vlanif30] quit

（3）配置ARP表项严格学习功能，使网关设备只对自己主动发送的ARP请求报文的应答报文触发本学习ARP表项，其他设备主动向网关设备发送的ARP报文不能触发本学习ARP表项。防止从伪造的ARP报文中学习ARP表项。

[HUAWEI] arp learning strict

（4）配置ARP表项固化模式为fixed-mac方式。使网关设备对收到的ARP报文中的MAC地址与ARP表中对应表项的MAC地址进行匹配检查，直接丢弃MAC地址不匹配的ARP报文。

[HUAWEI] arp anti-attack entry-check fixed-mac enable

（5）配置免费ARP报文主动丢弃功能。使网关设备直接丢弃免费ARP报文。

[HUAWEI] arp anti-attack gratuitous-arp drop

（6）配置根据源 IP地址进行ARP Miss消息限速，对Server（IP地址为10.10.10.2）的ARP Miss消息进行限速，允许Switch每秒最多处理该 IP地址触发的40个ARP Miss消息；对于其他用户，允许Switch每秒最多处理同一个源 IP地址触发的20个ARP Miss消息。

[HUAWEI] arp-miss speed-limit source-ip maximum20

[HUAWEI] arp-miss speed-limit source-ip 10.10.10.2 maximum40

（7）配置基于接口的ARP表项限制，使GE1/0/1接口最多可以学习到20个动态ARP表项。

[HUAWEI] interface gigabitethernet 1/0/1

[HUAWEI-GigabitEthernet1/0/1] arp-limit vlan 10 maximum20

[HUAWEI-GigabitEthernet1/0/1] quit

（8）配置根据源MAC地址进行ARP限速，对用户User1（MAC地址为1–1–1）进行ARP报文限速，每秒最多只允许10个该MAC地址的ARP报文通过。

[HUAWEI] arp speed-limit source-mac1-1-1maximum 10

（9）配置根据源IP地址进行ARP限速，对用户User3（IP地址为9.9.9.2）进行ARP报文限速，每秒最多只允许10个该IP地址的ARP报文通过。

[HUAWEI] arp speed-limit source-ip 9.9.9.2 maximum10

配置好后，可使用display arp learning strict命令查看全局已经配置ARP表项严格学习功能，以验证配置结果。具体如下。

```
[HUAWEI] display arp learning strict
The global configuration:arp learning strict
Interface                LearningStrictState
------------------------------------------------------------
------------------------------------------------------------
Total:0
Force-enable:0
Force-disable:0
```

还可通过display arp-limit命令查看接口可以学习到的动态ARP表项数目的最大值；通过display arp anti-attack configuration all命令查看当前ARP防攻击配置情况。 可通过display arp packet statistics命令查看ARP处理的报文统计数据，具体如下。

```
[HUAWEI] display arp packet statistics
ARP Pkt Received: sum    8678904
ARP-Miss Msg Received: sum    183
ARP Learnt Count:    sum    37
ARP Pkt Discard For Limit:    sum    146
ARP Pkt Discard For SpeedLimit:    sum    40529
ARP Pkt Discard For Proxy Suppress:    sum    0
ARP Pkt Discard For Other:    sum    8367601
ARP-Miss Msg Discard For SpeedLimit:    sum    20
ARP-Miss Msg Discard For Other:    sum    104
```

由显示信息可知，Switch上产生了ARP报文和ARP Miss消息丢弃计数，表明ARP安全功能已经生效。

16.5.2 防止ARP中间人攻击配置示例

本示例拓扑结构如图 16-2所示，SwitchA通过GE2/0/1接口连接DHCP Server，通过GE1/0/1和GE1/0/2接口分别连接DHCP客户端UserA和UserB，通过GE1/0/3接口 连接静态配置 IP 地址的用户 UserC。SwitchA的GE1/0/1、GE1/0/2、GE1/0/3、GE2/0/1 接口都属于 VLAN10。现希望能够防止 ARP 中间人攻击，避免合法用户的数据被中间人窃取，同时希望能够了解当前ARP中间人攻击的频率和范围。

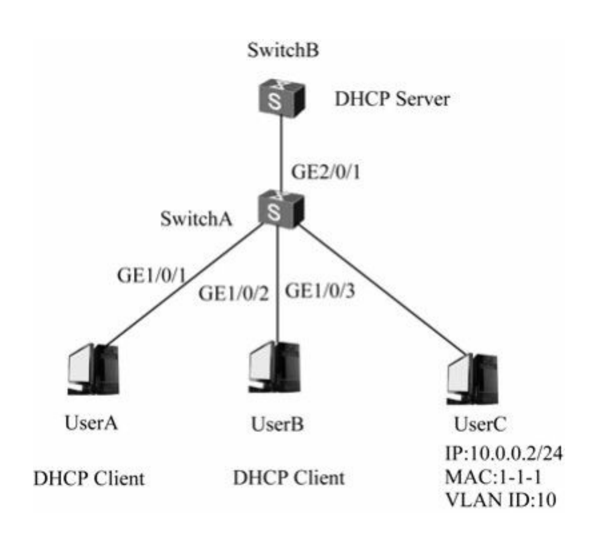


图16-2 防止ARP中间人攻击配置示例拓扑结构

1．基本配置思路分析

本示例可以采取以下方法来预防ARP中间人攻击。

（1）使能动态 ARP 检测功能，使SwitchA 对收到的 ARP 报文对应的源 IP地址、源MAC地址、VLAN以及接口信 息进行DHCP Snooping绑定表匹配检查，防止ARP中间人攻击。

（2）使能动态ARP检测丢弃报文告警功能，使SwitchA开始统计丢弃的不匹配DHCP Snooping 绑定表的 ARP 报文数量，并在丢弃数量超过告警阈值时能以告警的方式提醒管理员，这样可以使管理员根据告警信息以及报文丢弃计数来了解当前ARP中间人攻击的频率和范围。

（3）配置DHCP Snooping功能，并为UserC配置静态绑定表（对于采用DHCP自动分配 IP地址的UserA和UserB，在设备使能DHCP Snooping功能后，当他们上线时设备会自动生成DHCP Snooping绑定表），使动态ARP检测功能生效。

2．具体配置步骤

（1）创建VLAN10，并将GE1/0/1、GE1/0/2、GE1/0/3、GE2/0/1接口加入VLAN10中。

```
<HUAWEI>system-view
[HUAWEI] sysname SwitchA
[SwitchA] vlan batch 10
[SwitchA] interface gigabitethernet 1/0/1
[SwitchA-GigabitEthernet1/0/1] port link-type access
[SwitchA-GigabitEthernet1/0/1] port default vlan 10
[SwitchA-GigabitEthernet1/0/1] quit
[SwitchA] interface gigabitethernet 1/0/2
[SwitchA-GigabitEthernet1/0/2] port link-type access
[SwitchA-GigabitEthernet1/0/2] port default vlan 10
[SwitchA-GigabitEthernet1/0/2] quit
[SwitchA] interface gigabitethernet 1/0/3
[SwitchA-GigabitEthernet1/0/3] port link-type access
[SwitchA-GigabitEthernet1/0/3] port default vlan 10
[SwitchA-GigabitEthernet1/0/3] quit
```

[SwitchA] interfacegigabitethernet 2/0/1

[SwitchA-GigabitEthernet2/0/1] port link-type trunk

[SwitchA-GigabitEthernet2/0/1] port trunk allow-pass vlan 10

[SwitchA-GigabitEthernet2/0/1] quit

（2）使能动态 ARP 检测功能和动态 ARP 检测丢弃报文告警功能。在用户侧的GE1/0/1、GE1/0/2、GE1/0/3接口下使能动态ARP检测功能和动态ARP检测丢弃报文告警功能。因为这3个接口的配置完全一样，所以下面仅以GE1/0/1接口为例进行介绍。

[SwitchA] interfacegigabitethernet 1/0/1

[SwitchA-GigabitEthernet1/0/1] arp anti-attack check user-bind enable

[SwitchA-GigabitEthernet1/0/1] arp anti-attack check user-bind alarm enable

[SwitchA-GigabitEthernet1/0/1] quit

（3）配置DHCP Snooping功能。

[SwitchA] dhcp enable

[SwitchA] dhcp snooping enable #---全局使能DHCP Snooping功能

[SwitchA] vlan 10

[SwitchA-vlan10] dhcp snooping enable #---在VLAN10内使能DHCP Snooping功能，这就会为VLAN 10中的动态IP地址用户UserA和UserB自动生成绑定表

[SwitchA-vlan10] quit

[SwitchA] interfacegigabitethernet 2/0/1

[SwitchA-GigabitEthernet2/0/1] dhcp snooping trusted #---配置接口GE2/0/1为DHCP Snooping信任接口，所有接口缺省均为非信任端口

[SwitchA-GigabitEthernet2/0/1] quit

[SwitchA] user-bind static ip-address 10.0.0.2 mac-address 0001-0001-0001 interface gigabitethernet 1/0/3 vlan 10 #---在信任接口GE2/0/1上为采用静态IP地址分配的UserC用户配置静态绑定表

配置好后，可使用display arp anti-attack configuration check user-bind interface命令查看各接口下动态ARP检测的配置信息，以下是GE1/0/1接口上的动态ARP检测的配置信息。可以看到已使能了动态ARP检测功能和动态ARP检测丢弃报文告警功能。

[SwitchA] display arp anti-attack configuration check user-bind interfacegigabitethernet1/0/1

arp anti-attack check user-bind enable

arp anti-attack check user-bind alarm enable

还可通过display arp anti-attack statistics check user-bind interface命令查看各接口下动态ARP检测的ARP报文丢弃计数，以下是GE1/0/1接口下动态ARP检测的ARP报文丢弃计数。

[SwitchA] display arp anti-attack statistics check user-bind interfacegigabitethernet 1/0/1

Dropped ARP packet number is 966

Dropped ARP packet number since the latest warning is 605

由显示信息可知，GE1/0/1接口下产生了ARP报文丢弃计数和丢弃的ARP报文告警数，表明防ARP中间人攻击功能已经生效。当在各接口下多次执行命令 display arp anti-attack statistics check user-bind interface时，我们就可根据显示信息中“Dropped ARP packet number is”字段值的变化来了解ARP中间人攻击频率和范围。

第17章　AAA配置与管理

AAA是Authentication（认证）、Authorization（授权）和Accounting（计费）的简称，是网络安全的一种管理机制，提供了认证、授权、计费3种安全功能。同时提供本地认证/授权方式、RADIUS服务器认证/授权和计费方式、HWTACACS服务器认证/授权和计费三种AAA方案。后面两种可视为“委托认证/授权/计费”方式，因为这两种方式中的认证/授权/计费功能的实现不是由本地设备完成的，而是所配置的远程RADIUS服务器或HWTACACS服务器完成的。

与下一章介绍的NAC方案中的802.1x认证、MAC地址和Portal认证方式基于接入设备接口（仅可在接入设备上部署）进行的认证方式不同，AAA采用基于用户（可以是所有用户，也可以是特定用户组中的用户）进行认证、授权和计费的方案。当然，在NAC的各种认证方式中，也是需要借助AAA方案中配置的本地用户信息或者远程RADIUS服务器上配置的用户信息进行认证，所以本章的内容也是下章内容的基础。

本章将全面介绍华为S系列交换机所支持的以上三种AAA方案配置与管理方法（这3种方案的许多配置是完全相同或相近的，所以整体来说本章内容并不复杂）。不过，在学习本章内容前，建议先学习一种或多种RADIUS服务器或者HWTACACS服务器的配置方法，因为在后面两种委托认证/授权/计费方案中用户属性及授权属性均是在对应的服务器上配置的。目前这两种服务器厂商基本上都有相应的服务器软件，Windows和Linux服务器操作系统中也有提供RADIUS服务器功能。

17.1 AAA基础

AAA 是 Authentication（认证）、Authorization（授权）和 Accounting（计费）的简称，提供了认证、授权、计费3种安全功能。其中“认证”是用来验证用户是否可以获得网络访问权；“授权”是授权通过认证的用户可以使用哪些服务；“计费”是记录通过认证的用户使用网络资源的情况。当然，在实际网络应用中，可以只使用 AAA 提供的一种或两种安全服务。例如，公司仅仅想让员工在访问某些特定资源的时候进行身份认证，如果还希望对员工使用网络的情况进行记录，那么还需要配置计费服务器。

17.1.1 AAA的基本构架

AAA是采用“客户端/服务器”（C/S）结构，其中AAA客户端（也称网络接入服务器——NAS）就是使能了 AAA 功能的网络设备（可以是网络中任意一台设备，不一定是接入设备，而且可以在网络中多个设备上使能），而AAA服务器就是专门用来认证、授权和计费的服务器（可以由服务器主机配置，也可以由提供了对应服务器功能的网络设备上配置），如图17-1所示。



图17-1 AAA的基本构架图

在设备上使能了AAA功能后，当用户需要通过AAA客户端访问某个网络前，需要先从AAA服务器中获得访问该网络的权限。但这个任务通常不是由担当AAA客户端的设备自己来完成的，而是通过设备把用户的认证、授权、计费信息发送给 AAA 服务器来完成的。当然，如果在担当AAA客户端的设备上同时配置了相应的AAA服务器功能，则此时客户端和服务器端就为一体了，这时实现的是 AAA 本地认证和授权（本地方式不提供计费功能 ）了。

1．AAA认证

华为S系列交换机的AAA功能支持以下认证方式。

（1）不认证：对用户非常信任，不对其进行合法检查，一般情况下不采用这种方式。

（2）本地认证：将用户信息配置在本地设备上。本地认证的优点是速度快，可以为运营商降低成本，缺点是存储信息量受设备硬件条件限制。

（3）远程认证：将用户信息配置在AAA认证服务器上。支持通过RADIUS（Remote Authentication Dial In User Service，远程认证拨入用户服务）协议或HWTACACS（HuaWei Terminal Access Controller Access Control System，华为终端访问控制系统）协议进行远程认证。

2．AAA授权

华为S系列交换机的AAA功能支持以下5种授权方式。

（1）不授权：不对用户进行授权处理。

（2）本地授权：根据本地设备为本地用户账号配置的相关属性（如允许使用的接入服务类型和FTP访问目录等）进行授权。

（3）HWTACACS授权：由HWTACACS服务器对用户进行远程授权。

（4）if-authenticated 授权：如果用户通过了认证，而且使用的认证模式是本地或远程认证，则直接为用户授权。

（5）RADIUS认证成功后授权：**RADIUS**协议的认证和授权是绑定在一起的，不能单独使用**RADIUS**进行授权。

3．计费

华为S系列交换机的AAA功能支持以下3种计费方式（不支持本地计费方式 ）。

（1）不计费：不对用户计费。

（2）RADIUS计费：设备将计费报文送往RADIUS服务器，由RADIUS服务器完成对用户的计费。

（3）HWTACACS计费：设备将计费报文送往HWTACACS服务器，由HWTACACS服务器完成对用户的计费。

17.1.2 AAA基于域的用户管理

华为S系列交换机通过域来进行AAA用户管理，每个域下可以应用不同的认证、授权和计费方案，以及RADIUS或者HWTACACS服务器模板，相当于对用户进行分类管理。属于域中的用户通过在该域中应用的认证、授权和计费方案进行认证、授权和计费，所以本章后面将要介绍的 AAA 方案配置中，一定要在对应的域下被绑定、应用才能对具体用户生效。

缺省情况下，设备存在配置名为 default 和 default_admin 两个域，全局缺省普通域为default，全局缺省管理域为default_admin。两个域均不能删除，只能修改。当无法确认接入用户的域时使用缺省域， default域为接入用户的缺省域，缺省为本地认证；default_admin域为管理员账户（如http、SSH、telnet、terminal和

ftp用户）的缺省域，缺省为本地认证。

用户所属的域是由域分隔符后的字符串来决定的。域分隔符可以是为“@”、“|”、“%”等符号，如user@huawei就表示属于huawei域。如果用户名中没有带@，就属于系统缺省的default域。

说明

自定义的域可以同时被配置成全局缺省普通域和全局缺省管理域。但域下配置的 授权信息较AAA服务器的授权信息优先级低，即优先使用AAA服务器下发的授权属性，在 AAA 服务器无该项授权或不支持该项授权时域的授权属性才生效。当然通常是两者配置的授权属性一致。

17.1.3 RADIUS协议

RADIUS 最初仅是针对拨号用户的 AAA 协议，后来随着用户接入方式的多样化发展，RADIUS也适应多种用户接入方式，如以太网接入、ADSL接入。它通过认证授权来提供接入服务，通过计费来收集、记录用户对网络资源的使用。该协议定义了基于UDP 的RADIUS帧格式及其消息传输机制，并规定UDP端口1812、1813分别作为认证（包括授权）、计费端口。

1．RADIUS服务器

RADIUS服务器程序一般运行在中心计算机或工作站上，维护相关的用户认证和网络服务访问信息，负责接收用户连接请求并认证用户，然后给客户端返回所有需要的信息（如接受/拒绝认证请求）。RADIUS服务器通常要维护以下3个数据库。

（1）Users：用于存储用户信息（如用户名、口令以及使用的协议、IP地址等配置信息）。

（2）Clients：用于存储RADIUS客户端的信息（如接入设备的共享密钥、IP地址等）。

（3）Dictionary：用于存储RADIUS协议中的属性和属性值含义的信息。

2．RADIUS客户端

RADIUS客户端程序一般位于网络接入服务器NAS（Network Access Server）设备上，可以遍布整个网络，负责传输各个接入网络用户信息到指定的RADIUS服务器，然后根据从RADIUS服务器返回的信息进行相应处理（如接受/拒绝用户接入）。

3．安全机制

RADIUS客户端和RADIUS服务器之间认证消息的交互是通过共享密钥来对传输数据加密的，但共享密钥不通过网络来传输，增强了信息交互的安全性。

4．认证和计费消息流程

RADIUS客户端与服务器间的信息交互流程如图17-2所示。

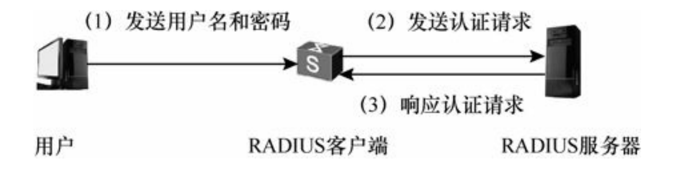


图17-2 RADIUS客户端与服务器间的信息交互流程

（1）用户访问RADIUS客户端设备时，会按照提示输入用户名和密码，发送给客户设备。

（2）客户端设备在收到用户发来的用户名和密码信息向 RADIUS 服务器发送认证请求。

（3）RADIUS服务器接收到合法的请求后，完成认证，并把所需的用户授权信息返回给接入设备；对于非法的请求，RADIUS服务器返回认证失败的信息给客户端设备。

RADIUS计费的信息交互流程和认证/授权的信息交互流程类似。

17.1.4 HWTACACS协议

HWTACACS是在TACACS（RFC 1492）基础上进行了功能增强的安全协议。该协议与RADIUS协议类似，也是采用C/S模式实现NAS与HWTACACS服务器之间的通信。

HWTACACS 协议主要用于点对点协议 PPP 和 VPDN（Virtual Private Dial-up Network，虚拟私有拨号网络）接入用户及终端用户的认证、授权和计费。其典型应用是对需要登录到设备上进行操作的终端用户进行认证、授权和计费。同样，这时的设备是作为 HWTACACS的客户端，负责将用户名和密码发给HWTACACS服务器进行验证。

HWTACACS协议与RADIUS协议都实现了认证、授权、计费的功能，它们有很多相似点：结构上都采用 C/S 模式，都使用公共密钥对传输的用户信息进行加密。但与RADIUS 相比，HWTACACS 具有更加可靠的传输和加密特性，更加适合于安全控制。HWTACACS协议与RADIUS协议的主要区别如表17-1所示。

表17-1 HWTACACS协议与RADIUS协议的比较

HWTACACS	RADIUS
使用 TCP，网络传输更可靠	使用 UDP
除了标准的 HWTACACS 报文头，对报文主体全部进行加密	只是对认证报文中的密码字段进行加密
认证与授权分离	认证与授权一起处理
适于进行安全控制	适于进行计费
支持对设备上的配置命令进行授权使用	不支持对设备上的配置命令进行授权

说明

HWTACACS协议与其他厂商支持的TACACS+协议都实现了认证、授权、计费的功能。而且，HWTACACS和TACACS+的认证流程与实现方式是一致的，HWTACACS协议能够完全兼容TACACS+协议。

17.1.5 AAA特性的产品支持

华为S系列交换机除了支持通过RADIUS协议或HWTACACS协议进行认证、授权、计费外，还支持本地认证和授权。且理论上，设备支持本地、RADIUS、HWTACACS三种协议间的认证、授权、计费的随意组合，比如本地认证、本地授权和RADIUS计费。但是在实际应用中，常见的是三种协议的单独应用。

1．本地认证、授权

如果需要对用户进行认证或授权，但是在网络中没有部署 RADIUS 服务器和HWTACACS 服务器，那么可以采用本地方式进行认证和授权。本地方式进行认证和授权的优点是速度快（因为直接在本地设备上进行处理），可以降低运营成本；缺点是存储信息量受设备硬件条件限制。通常仅对管理员用户采用本地方式进行认证和授权，毕竟管理员用户数不会很多。

2．RADIUS认证、计费

因为RADIUS服务器不是位于设备上，而是位远程主机上，且设备与RADIUS服务器之间的通信是加密的，所以采用RADIUS方式进行认证、计费可以防止非法用户对网络的攻击，常应用在既要求较高安全性又要求控制远程用户访问权限的网络环境中。但**RADIUS**服务器不支持单独授权功能，必须与认证功能一起，只要使能它的认证功能，就同时使能了它的授权功能。

3．HWTACACS认证、授权、计费

与RADIUS服务器一样，采用HWTACACS方式进行认证、授权、计费也可以防止 非法用户对网络的攻

击，还支持为用户进行具体的命令行授权。且与 RADIUS 相比， HWTACACS 的认证、授权和计费是单独进行的，可以单独配置和使能，在一些大的网络中，更加方便部署多台用途不同的HWTACACS服务器。

此外，设备还支持一个方案中使用多种协议模式，比如本地认证还常用于RADIUS认证和HWTACACS认证的备份认证方案，本地授权作为HWTACACS授权的备份授权方案。

配置以上三种AAA方案的流程如图17-3所示，它们的总体配置流程基本一样，只是用户认证信息的配置不一样而已。下面介绍的各 AAA 方案的具体配置步骤也是按照这个流程进行的。

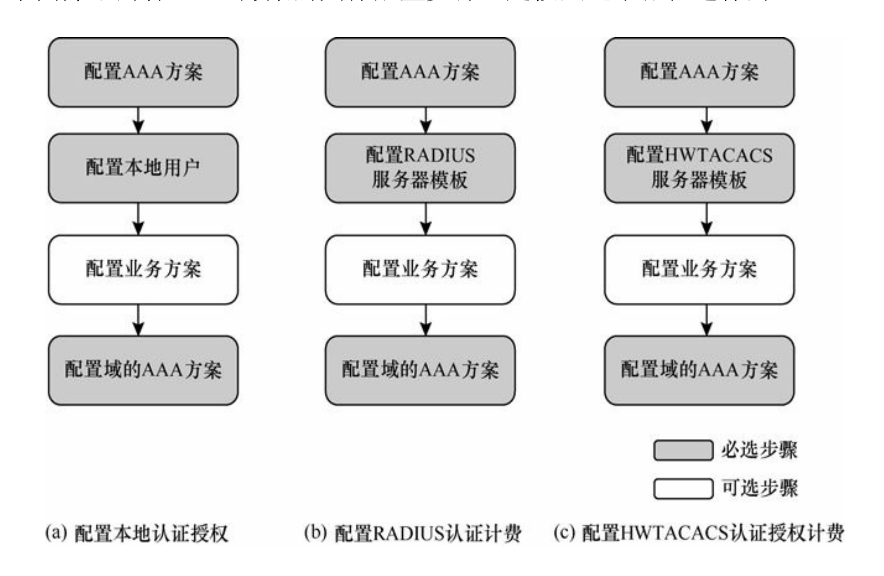


图17-3 三种AAA方案的配置流程

17.2 本地方式认证和授权配置

本地方式就是把S系列交换机同时配置为AAA客户端和AAA认证、授权服务器（不能另外配置计费功能，总是采用不计费方案），认证和授权信息是在本地设备上配置的，无需另外配置专门的认证服务器。配置采用本地方式进行认证和授权后，设备根据本地的用户信息对接入用户进行认证和授权。本地方式进行认证和授权的优点是速度快，可以降低运营成本，缺点是存储信息量受设备硬件条件限制。

根据图 17-3（a）可以得出本地认证、授权配置流程为配置 AAA 方案→配置本地用户→配置业务方案→配置域的AAA方案。其中第三步的“配置业务方案”是可选的，通常是采用缺省配置，仅当需要特定的 IP 业务调用本地认证、授权方案时才需要配置，其余三项是必选的，但前面两项配置任务没有严格的先后次序，都是为最后一项配置任务“配置域的AAA方案”而服务的。下面分别介绍这4项配置任务的具体配置步骤。

17.2.1 配置AAA方案

配置AAA方案就是配置AAA中的认证、授权和计费方案，用于后面将要介绍的“域的AAA 方案”中绑定这些方案使用（所配置的各种方案只有在域中绑定后才能得到应用）。在本地方式中仅支持认证和授权方案，所以仅需要在认证方案中配置认证模式为本地认证，在授权方案中配置授权模式为本地授权，其他配置均为可选的，具体配置步骤如表17-2所示。

表17-2 本地认证、授权中的AAA方案配置步骤

<table>
<tr><th>步骤</th><th>命令</th><th>说明</th></tr>
<tr><td>1</td><td>system-view
例如：<HUAWEI> system-view</td><td>进入系统视图</td></tr>
<tr><td>2</td><td>aaa
例如：[HUAWEI] aaa</td><td>进入 AAA 视图</td></tr>
<tr><td colspan="3">配置 AAA 认证方案</td></tr>
<tr><td>3</td><td>authentication-scheme authentication-scheme-name
例如：[HUAWEI-aaa] authentication-scheme scheme0</td><td>创建一个认证方案，并进入认证方案视图或直接进入一个已存在的认证方案视图。参数 scheme-name 用来指定认证方案名，为 1～32 个字符，不支持空格，不区分大小写，且不能包含以下字符：“\”、“/”、“:”、“<”、“>”、“|”、“@”、“'”、“%”、“*”、“"”、“?”。不同系列交换机所支持的认证方案数不一样，S2700 系列最多只支持 16 个，而 S3700/5700/6700/7700//9700 系列最多支持 32 个，S9300/9300E 最多支持 128 个
缺省情况下，设备中有一个名为“default”的认证方案，不能删除，只能修改，可用 undo authentication-scheme authentication-scheme-name 命令删除指定的认证方案
【说明】“default”认证方案的策略为（1）认证模式采用本地认证；（2）用户提升级别模式采用本地认证；（3）认证失败则强制用户下线</td></tr>
<tr><td>4</td><td>authentication-mode local
例如：[HUAWEI-aaa-authen-scheme0] authentication-mode local</td><td>（可选）配置认证模式为本地认证
因为 S 系列交换机在缺省情况下，认证模式就为本地认证，所以缺省情况下可不用配置本命令，如果已配置为其他认证模式，则也可用 undo authentication-mode 命令恢复当前认证方案使用的认证模式为缺省的本地认证模式</td></tr>
<tr><td>5</td><td>authentication-super { hwtacacs | radius | super }* [none]
例如：[HUAWEI-aaa-authen-scheme0] authentication-super radius</td><td>（可选）配置当前认证模板对用户提升级别进行认证时采用的认证模式。命令中的选项说明如下。
（1）hwtacacs：可多选项，指定采用 HWTACACS 模式对用户级别提升进行认证
（2）radius：可多选项，指定采用 RADIUS 模式对用户级别提升进行认证
（3）super：可多选项，指定采用本地认证的模式对用户级别提升进行认证
（4）none：可选项，指定无需进行认证，即直接让用户更改用户级别
【说明】可以同时配置多种认证模式。此时认证模式的执行顺序为配置的先后顺序。只有在当前认证模式没有响应（不是认证失败）的情况下，设备才会采用下一种认证模式；如果用户没有通过当前认证模式的认证，则设备不会再跳转到下一个认证模式对用户进行认证
缺省情况下，用户级别提升时认证模式为本地模式，可用 undo authentication-super 命令恢复用户级别提升时采用的认证模式为缺省情况</td></tr>
</table>

（续表）

步骤	命令	说明
6	**quit** 例如：[HUAWEI-aaa-authen-scheme0] **quit**	退出认证方案视图，返回 AAA 视图
7	**domainname-parse-direction** { **left-to-right** | **right-to-left** } 例如：[HUAWEI-aaa] **domainname-parse-direction left-to-right**	（可选）配置用户名和域名解析的方向。命令中的选项说明如下。 （1）**left-to-right**：二选一选项，指定域名解析方向为从左向右 （2）**right-to-left**：二选一选项，指定域名解析方向为从右向左 【说明】这是为了便于系统识别域用户名格式，但必须与在 AAA 视图下使用 **domain-location** { **after-delimiter** | **before-delimiter** } 命令配置的域用户名格式一致，其中选择 **after-delimiter** 选项时，表示指定域名在分隔符后；选择 **before-delimiter** 选项时，表示指定域名在分隔符前 域用户名通常采用"纯用户名@域名"格式，即@符号后面的部分为域名。如果配置的是 **domain-location after-delimiter** 命令，指定域名在分隔符后。现假设域用户名为 username@dom1 @dom2，如果采用从左向右解析，则用户名为 username，域名为 dom1 @dom2，相反，如果采用从右向左解析，则用户名为 username@dom1，域名为 dom2 如果配置的是 **domain-location before-delimiter** 命令，指定域名在分隔符前，如果采用从左向右解析时，则用户名为 dom1@dom2，域名为 username；如果采用从右向左解析，则用户名为 dom2，域名为 username@dom1 缺省情况下，域名解析方向为从左向右，可用 **undo domainname-parse-direction** 命令恢复域名解析方向为缺省设置
	配置 AAA 授权方案	
8	**authorization-scheme** *authorization-scheme-name* 例如：[HUAWEI-aaa] **authorization-scheme** scheme0	创建一个授权方案，并进入授权方案视图或直接进入一个已存在的授权方案视图。其他说明与前面第 3 步介绍的认证方案是完全一样的（不同的只是这里是授权方案），参见即可 缺省情况下，设备中有一个名为"default"的授权方案，不能删除，只能修改，可用 **undo authorization-scheme** *authorization-scheme-name* 命令删除指定的授权方案
9	**authorization-mode local** [**none**] 例如：[HUAWEI-aaa-author-scheme0] **authorization-mode local**	配置本地授权模式。如果同时选择了 **none** 可选项，则表示无需授权
10	**quit** 例如：[HUAWEI-aaa-author-scheme0] **quit**	退出授权方案视图，返回 AAA 视图
11	**authorization-modify mode** { **modify** | **overlay** } 例如：[HUAWEI-aaa] **authorization-modify mode modify**	（可选）配置授权服务器下发的用户授权信息的生效模式。授权服务器可向设备单独，或同时下发 ACL 规则、动态 VLAN 等用户授权信息。命令中的选项说明如下。 （1）**modify**：二选一选项，指定授权服务器下发的用户授权信息的生效模式为修改模式，新下发的授权信息仅按对应下发的属性类别覆盖上一次下发的授权信息 （2）**overlay**：二选一选项，指定授权服务器下发的用户授权信息的生效模式为覆盖模式，新下发的授权信息覆盖上一次下发的所有属性类别的授权信息 缺省情况下，设备上用户授权信息的生效模式为覆盖模式，即新下发的用户授权信息将会覆盖前次下发的所有的用户授权信息，可用 **undo authorization-modify mode** 命令恢复授权服务器下发的用户授权信息的生效模式为覆盖模式

17.2.2 配置本地用户

当采用本地方式进行认证和授权时，需要在本地设备上配置用户的认证和授权信息，如用于认证的用户名和密码，用于授权的用户优先级、用户组、允许接入的服务类型、可建立的连接数和FTP访问目录等。但这些均需要在AAA视图下进行配置，具体的配置步骤如表17-3所示（除创建本地用户命令外，其他配置均为可选的，均有对应的缺省配置 ）。

表17-3 本地认证、授权中的本地用户配置步骤

步骤	命令	说明
1	**system-view** 例如：<HUAWEI> **system-view**	进入系统视图
2	**aaa** 例如：[HUAWEI] **aaa**	进入 AAA 视图
3	**local-user** *user-name* **password cipher** *password* 例如：[HUAWEI-aaa] **local-user** user1@mydomain **password cipher** admin	创建本地用户账号，并配置本地用户账号的登录密码。命令中的参数说明如下： （1）*user-name*：指定本地用户的用户名，为 1～64 个字符，**不支持空格，不区分大小写**。如果用户名中带域名分隔符（如“@”、“|”、“%”等符号），则缺省情况下认为分隔符前面的部分是用户名，后面部分是域名。如果没有分隔符，则整个字符串为用户名，缺省采用 **default** 域认证 （2）**cipher** *password*：指定本地用户登录密码，以密文显示。可以是 1～16 个字符的明文密码，**不支持空格、单引号和问号，区分大小写**；也可是 32 位密文密码 缺省情况下，没有创建本地用户，可用 **undo local-user** *user-name* 删除指定的本地用户账户
4	**local-user** *user-name* **privilege level** *level* 例如：[HUAWEI-aaa] **local-user** user1@mydomain **privilege level** 10	（可选）配置本地用户的级别。命令中的参数说明如下。 （1）*user-name*：指定要配置用户级别的本地用户名 （2）*level*：为由 *user-name* 参数指定的本地用户配置用户级别，取值范围为 0～15 的整数，且值越大，级别越高。不同级别的用户登录后，只能使用等于或低于自己级别的命令 缺省情况下，本地用户（如 Telnet 用户、SSH 用户）的优先级由管理模块（也就是按照系统缺省的 4 个命令级别来分类，具体参见本书第 2 章 2.1.4）来决定，可用 **undo local-user** *user-name* **privilege level** 命令将指定的本地用户的优先级恢复为缺省配置
5	**local-user** *user-name* **user-group** *user-group-name* 例如：[HUAWEI-aaa] **local-user** user1@mydomain **user-group** test1	（可选）配置本地用户加入指定用户组，**但 S2700/3700 系列不支持，在 S5700 系列中仅 S5700HI 和 S5710EI 子系列支持**。命令中的参数说明如下。 （1）*user-name*：指定要加入用户组的本地用户 （2）*user-group-name*：指定由 *user-name* 参数指定的用户要加入的用户组，为 1～64 个字符，区分大小写，不支持空格，可以设定为包含数字、字母和“*”、“#”等特殊字符的组合 【说明】系统对用户权限的管理是通过用户组实现的，用户加入相应的用户组就能获得相应的用户权限。用户组是在系统视图下使用 user-group group-name 命令创建的，在用户组下面可以配置许多授权信息，具体参见本书第 18 章中的 18.2.18 节介绍。一个用户组可被多个本地用户引用，但一个本地用户只能属于一个用户组。缺省用户组、被本地用户或在线用户引用的用户组不能删除 如果本地用户与对应域下面同时都配置了用户组，且所配置的用户组不同（相当于下发的授权信息不同），则只有本地用户配置的用户组生效 缺省情况下，本地用户不属于任何用户组，可用 **undo local-user** *user-name* **user-group** 命令取消指定本地用户加入用户组

（续表）

步骤	命令	说明
6	**local-user** *user-name* **idle-timeout** *minutes* [*seconds*] 例如：[HUAWEI-aaa] **local-user** user1@mydomain **idle-timeout** 1 30	（可选）配置指定用户的闲置切断时间（也就是配置闲置多长时间后把对应用户下线）。命令中的参数说明如下。 （1）*user-name*：指定要配置闲置切断时间的本地用户 （2）*minutes* [*seconds*]：指定由 *user-name* 参数指定的用户的闲置切断时间的分钟和秒数，取值范围分别为 0～35 791 的整数和 0～59 的整数。当这两个值均为 0 时表示关闭超时断连功能 缺省情况下，超时时间为 5min，可用 **undo local-user** *user-name* **idle-timeout** 命令恢复指定本地用户的断连超时时间为缺省值
7	**local-user** *user-name* **service-type** { **8021x** | **bind** | **ftp** | **http** | **ppp** | **ssh** | **telnet** | **terminal** | **web** | **x25-pad** }* 例如：[HUAWEI-aaa] **local-user** user1@mydomain **service-type 8021x**	（可选）配置允许本地用户的接入类型。命令中的参数和选项说明如下。 （1）*user-name*：指定要配置接入类型的本地用户 （2）**8021x**：可多选项，指定用户类型为 802.1x 用户 （3）**bind**：可多选项，指定用户类型为 IP 会话用户 （4）**ftp**：可多选项，指定用户类型为 FTP 用户 （5）**http**：可多选项，指定用户类型为 HTTP 用户 （6）**ppp**：可多选项，指定用户类型为 PPP 用户 （7）**ssh**：可多选项，指定用户类型为 SSH 用户 （8）**telnet**：可多选项，指定用户类型为 Telnet 用户 （9）**terminal**：可多选项，指定用户类型为 Console 口用户、TTY 用户 （10）**web**：可多选项，指定用户类型为 Web 认证用户 （11）**x25-pad**：可多选项，指定用户类型为 X25-PAD 用户 缺省情况下，本地用户可以使用所有的接入类型，可用 **undo local-user** *user-name* **service-type** 命令将指定的本地用户的接入类型恢复为缺省配置
8	**local-user** *user-name* **ftp-directory** *directory* 例如：[HUAWEI-aaa] **local-user** user1@mydomain **ftp-directory** flash:/ftp	（可选）配置允许 FTP 用户访问的 FTP 目录。**当设备作为 FTP 服务器时，必须配置允许 FTP 用户访问的 FTP 目录，否则 FTP 用户无法访问设备**。命令中的参数说明如下。 （1）*user-name*：指定要配置访问目录的本地 FTP 用户 （2）*directory*：指定由 *user-name* 参数指定的 FTP 用户可访问的目录，为 1～64 个字符，不支持空格，区分大小写 缺省情况下，允许 FTP 用户访问的 FTP 目录为空，可用 **undo local-user** *user-name* **ftp-directory** 命令删除指定本地用户配置的 FTP 访问目录
9	**local-user** *user-name* **state** { **active** | **block** } 例如：[HUAWEI-aaa] **local-user** user1@mydomain **state block**	（可选）配置本地用户的状态。命令中的参数和选项说明如下。 （1）*user-name*：指定要配置状态的本地用户 （2）**active**：二选一选项，指定由 *user-name* 参数指定的本地用户为激活状态，接收该用户的认证请求并做进一步处理 （3）**block**：二选一选项，指定由 *user-name* 参数指定的本地用户为阻塞状态，拒绝该用户的认证请求 缺省情况下，本地用户的状态为激活态
10	**local-user** *user-name* **access-limit** *max-number* 例如：[HUAWEI-aaa] **local-user** user1@mydomain **access-limit** 30	（可选）配置指定用户可建立的最大连接数目。命令中的参数说明如下。 （1）*user-name*：指定要配置可建立的最大连接数目的本地用户 （2）*max-number*：指定由 *user-name* 参数指定的本地用户可建立的最大连接数目，取值范围不同系列有所不同，具体参见相应产品手册说明 缺省情况下，不限制用户可建立的连接数目，可用 **undo local-user** *user-name* **access-limit** 命令恢复为缺省情况

（续表）

步骤	命令	说明
11	**local-aaa-user wrong-password retry-interval** *retry-interval* **retry-time** *retry-time* **block-time** *block-time* 例如：[HUAWEI-aaa] **local-aaa-user wrong-password retry-interval** 10 **retry-time** 3 **block-time** 30	（可选）使能本地账号锁定功能，并配置用户的重试时间间隔、连续认证失败的限制次数及账号锁定时间。命令中的参数说明如下。 （1）*retry-interval*：指定本地用户每次重试的时间间隔，取值范围为 5～65 535 的整数分钟，超过这个时间该账户被锁定 （2）*retry-time*：指定本地用户连续认证失败的最大次数，取值范围为 3～65 535 的整数，超过这个次数即该账户被锁定。但这里的"认证失败次数"仅针对密码错误，其他本地认证错误不计入计数 （3）*block-time*：指定本地用户被锁定的时间，取值范围为 5～65 535 的整数分钟 缺省情况下，未使能本地账号锁定功能，可用 **undo local-aaa-user wrong-password** 命令去使能本地账号锁定功能
12	**return** 例如：[HUAWEI-aaa] **return**	退出 AAA 视图，直接返回用户视图
13	**local-user change-password** 例如：<HUAWEI> **local-user change-password**	（可选）修改本地用户的登录密码。为了保证低级别管理用户的密码安全，管理用户认证通过后，可以在用户视图下通过此命令来修改自己的登录密码。 本地认证通过的用户才可以执行该命令修改自己的密码。本地用户成功修改自己的密码后，下次登录需要输入新密码才可以认证通过。该命令本身是修改本地用户密码，不直接保存配置，但会以 **local-user password** 的命令形式保存修改结果。若系统等待超过 30s 后，用户还未输入用户名或新密码、确认密码时，密码修改将中断。用户输入 Ctrl+C 取消本次密码修改时，密码修改将中断。

【示例 】在本地认证中，通过 local-user change-password命令用户修改自己的密码。

```
<HUAWEI> local-user change-password
Please enter old password:
Please enter new password:
Please confirm new password:
Info: The password is changed successfully.
```

17.2.3 （可选）配置业务方案

"业务方案"其实也是一种授权方案，它是专门针对一些IP业务（如管理员权限、DHCP 服务、DNS 服务和策略路由等）所进行的授权，所也可称为"业务授权方案"。可通过配置业务方案管理用户的业务授权信息，但在业务方案下通常只需要使用admin-user privilege level命令配置管理员用户的用户级别，其余命令在业务方案被其他特性（比如IPSec特性）调用时才需要配置。具体的配置步骤如表17-4所示。

表17-4 本地认证、授权中的业务方案配置步骤

步骤	命令	说明
1	**system-view** 例如：<HUAWEI> **system-view**	进入系统视图
2	**aaa** 例如：[HUAWEI] **aaa**	进入 AAA 视图

（续表）

步骤	命令	说明
3	**service-scheme** *service-scheme-name* 例如：[HUAWEI-aaa] **service-scheme** svcscheme1	创建一个业务方案，并进入业务方案视图或直接进入一个已存在的业务方案视图。其他说明与17.2.1节表17-2中第3步介绍的认证方案是完全一样的（不同的只是这里是授权方案），参见即可 缺省情况下，设备中没有配置业务方案，可用 **undo service-scheme** *service-scheme-name* 命令删除指定的业务方案
4	**admin-user privilege level** *level* 例如：[HUAWEI-aaa-service-svcscheme1] **admin-user privilege level**10	配置本地用户可以作为管理员登录设备，并设置这些本地用户在设备中的管理员级别，取值范围是0～15的整数 【说明】如果用户的认证方式为本地认证，管理员用户级别可以采用以下三种方式配置，优先级由上到下依次降低。 （1）使用 **local-user privilege level** 命令配置的本地用户级别 （2）使用 **admin-user privilege level** 命令在域下配置的管理员用户级别 （3）使用 **user privilege** 命令在VTY模式下配置的用户级别 如果用户的认证方式为远端认证，管理员用户级别可以采用以下三种方式配置，优先级由上到下依次降低。 （1）认证通过后，服务器下发到设备中的用户级别 （2）使用 **admin-user privilege level** 命令在域下配置的管理员用户级别 （3）使用 **user privilege** 命令在VTY模式下配置的用户级别 如果对用户同时配置了远端认证和本地认证，且配置顺序是先远端认证再本地认证，管理员用户级别可以采用以下4种方式配置，优先级由上到下依次降低。 （1）认证通过后，服务器下发到设备中的用户级别。 （2）使用 **local-user privilege level** 命令配置的本地用户级别。**本地用户级别只在远端认证服务器没有响应时启用**。如果配置了本地用户级别，远端服务器认证响应通过后但是没有下发用户级别，此时本地用户的级别不会生效 （3）使用 **admin-user privilege level** 命令在域下配置的用户级别 （4）使用 **user privilege** 命令在VTY模式下配置的用户级别 缺省情况下，用户的用户级别为16，表示是一个无效值，用户不能作为管理员登录设备，可用 **undo admin-user privilege level** 命令指定当前用户不能作为管理员登录设备，并恢复用户级别为缺省级别
5	**dhcp-server group** *group-name* 例如：[HUAWEI-aaa-service-svcscheme1] **dhcp-server group** group1	（可选）设置业务方案下使用的DHCP服务器组，参数 *group-name* 用来指定DHCP服务器组名，为1～32个字符，区分大小写，不支持空格，必须是已由 **dhcp-server group** *group-name* 系统视图命令配置好相应的DHCP服务器组。仅 **S7700/9300/9300E/9700** 系列支持 【说明】通常情况下，一台DHCP中继会同时代理多台DHCP服务器，并向用户提供IP地址分配服务。此时可使用 **dhcp server group** 命令定义DHCP服务器组来统一管理该DHCP中继代理的DHCP服务器，从而在指定的DHCP服务器组中为通过DHCP中继接入的用户分配IP地址 缺省情况下，业务方案下没有配置DHCP服务器组，可用 **undo dhcp-server group** 命令取消业务方案使用的DHCP服务器组

（续表）

步骤	命令	说明
6	**ip-pool** *pool-name* [**move-to** *new-position*] 例如：[HUAWEI-aaa-service-svcscheme1] **ip-pool** ippool1 **move-to** ippool2	（可选）配置业务方案下可用的 DHCP IP 地址池或者移动已配置的地址池的位置，仅 **S7700/9300/9300E/9700** 系列支持。命令中的参数说明如下。 （1）*pool-name*：指定 IP 地址池名，为 1～64 个字符，可由以下字符组成：字母（包括大写字母“A”～“Z”和小写字母“a”～“z”）、数字（“0”～“9”）、点号（“.”）、短线（“-”）和下划线（“_”）。该 IP 地址池必须在上一步配置的 DHCP 服务器组中的 DHCP 服务器中已创建 （2）**move-to** *new-position*：可选参数，指定移动业务方案下已配置的 IP 地址池的位置信息，取值范围与域下已配置的 IP 地址池数相关，为一个可用的 IP 地址池名称 缺省情况下，业务方案下没有配置 IP 地址池，可用 **undo ip-pool** [*pool-name*]命令删除所有或者指定业务方案下的 IP 地址池
7	**dns** *ip-address* [**secondary**] 例如：[HUAWEI-aaa-service-svcscheme1] **dns** 10.10.10.1	（可选）配置主用或者备用（选择 **secondary** 可选项时）DNS 服务器地址。 缺省情况下，业务方案下没有配置主用和备用的 DNS 主用服务器，可用命令 **undo dns** [*ip-address*]删除指定的 DNS 服务器
8	**policy-route** *next-hop-ip-address* [*vlan-id*] 例如：[HUAWEI-aaa-service-svcscheme1] **policy-route 20.1.1.1**	（可选）配置业务方案下用户的策略路由功能，仅 S7700/9300/9300E/9700 系列支持。命令中的参数说明如下。 • *next-hop-ip-address*：指定策略路由的下一跳 IP 地址 • *vlan-id*：可选参数，指定源路由的 VLAN ID 缺省情况下，业务方案下没有配置策略路由功能，可用 **undo policy-route** 命令取消业务方案下用户的策略路由功能

17.2.4 配置域的AAA方案

在17.2.1节创建的认证和授权方案，在17.2.3节配置的业务方案也只有在本节介绍的“域的 AAA 方案”中绑定后才能得到应用（在本地认证、授权方案中无需配置像后面将要介绍的 **RADIUS**、**HWTACACS** 服务器模板）。不同的域可以绑定不同的以上认证、授权、业务方案，以便实现灵活的用户接入控制。域的 AAA 方案的具体配置步骤如表17-5所示。

表17-5 本地认证、授权中域的AAA方案配置步骤

步骤	命令	说明
1	**system-view** 例如：<HUAWEI> **system-view**	进入系统视图
2	**aaa** 例如：[HUAWEI] **aaa**	进入 AAA 视图
3	**domain** *domain-name* 例如：[HUAWEI-aaa] **domain** mydomain	创建域并进入域视图或进入一个已存在的域视图。参数 *domain-name* 用来指定域名，为 1～64 个字符，不支持空格，区分大小写，且不能包含以下字符：“-”、“*”、“?”、“"” 缺省情况下，设备存在两个域：default 和 default_admin。default 用于普通接入用户的域，default_admin 用于管理员的域。可用 **undo domain** *domain-name* 命令删除指定域

（续表）

<table>
<tr><th>步骤</th><th>命令</th><th>说明</th></tr>
<tr><td>4</td><td>authentication-scheme authentication-scheme-name
例如：[HUAWEI-aaa-domain-mydomain] authentication-scheme scheme1</td><td>在以上域中绑定要使用的认证方案。这些认证方案就是在 17.2.1 节表 17-2 中创建并配置的，选择所需的方案即可</td></tr>
<tr><td>5</td><td>authorization-scheme authorization-scheme-name
例如：[HUAWEI-aaa-domain-mydomain] authorization-scheme author1</td><td>在以上域中绑定要使用的授权方案。这些授权方案就是在 17.2.1 节表 17-2 中创建并配置的，选择所需的方案即可</td></tr>
<tr><td>6</td><td>user-group group-name
例如：[HUAWEI-aaa-domain-mydomain] user-group ftp</td><td>（可选）配置以上域中要下发授权的用户组，即仅对应用户组中的本地用户才可使用该域中的认证、授权、计费（本地方式中不支持计费功能，其他方式支持）和业务方案。参数 group-name 为要在域中要下发授权的用户组，为 1～64 个字符，区分大小写，不支持空格，可以设定为包含数字、字母和“*”、“#”等特殊字符的组合。本地用户是通过 17.2.2 节表 17-3 中的第 5 步加入到对应的用户组的
缺省情况下，未配置对域下的用户下发用户组授权，所有本地用户均可作为本域中的用户，可用 undo user-group 命令取消对域下的用户下发用户组授权</td></tr>
<tr><td>7</td><td>service-scheme service-scheme-name
例如：[HUAWEI-aaa-domain-mydomain] service-scheme services1</td><td>（可选）在以上域中绑定要使用的业务方案。这些业务方案就是在 17.2.3 节表 17-4 中创建并配置的，选择所需的方案即可。如果没有业务方案，则不用进行本步设置</td></tr>
<tr><td>8</td><td>state { active | block }
例如：[HUAWEI-aaa-domain-mydomain] state active</td><td>（可选）配置域的状态：激活（选择 active 二选一选项时）或者阻塞选择 block 二选一选项时）。当域处于阻塞态时，属于该域的用户不能登录。缺省情况下，域创建后处于激活状态，可用 undo state 命令恢复域的状态为激活状态</td></tr>
<tr><td>9</td><td>quit
例如：[HUAWEI-aaa-domain-mydomain] quit</td><td>退出域视图，返回 AAA 视图</td></tr>
<tr><td>10</td><td>domain-name-delimiter delimiter
例如：[HUAWEI-aaa-domain-mydomain] domain-name-delimiter @</td><td>（可选）配置域名分隔符，可以是 \/:<>|@'% 中的一个
缺省情况下，域名分隔符为@，可用 domain-name-delimiter 命令恢复域名分隔符为缺省的@</td></tr>
</table>

17.3 RADIUS方式认证、授权和计费配置

在前面介绍的本地方式认证、授权中使用的是在本地设备上储存的用户信息和属性进行认证和授权，这仅对于小型网络有效，因为设备上可以储存的用户数是很有限的。 所以，通常是采用像RADIUS服务器，或者HWTACACS服务器进行远程认证、授权，并且可以计费。而RADIUS服务器，或者HWTACACS服务器的AAA方案更加安全，可以避免设备遭受攻击。

本节要介绍使用RADIUS协议对接入用户进行认证、授权和计费的配置方法（不包括**RADIUS**服务器中用户账户信息和用户授权属性等自身的配置 ）。但RADIUS中的认证和授权是同步进行的，只要使能了其认证功能，也就同时使能了其授权功能。

说明

根据图17-3（b）可以得知RADIUS认证、授权和计费方式的配置流程。在RADIUS方式的配置中，与本地方式的配置中主要不同在于前面两项配置任务，本节仅介绍这两项配置任务的具体配置方法和步骤。

第三项的“配置业务方案”和与前面17.2.3节介绍的本地方式中的“配置业务方案”的配置方法和步骤完全一样，不同的只是这里配置的业务方案适用于RADIUS服务器认证、授权和计费。

第四项的“配置域的AAA方案”和17.2.4节介绍的本地方式中的“配置域的AAA方案”的配置方法与步骤总体也一样，有如下两处不同的地方。

（1）在17.2.4节表17-5中第5步中要用accounting-scheme accounting-scheme-name命令来替换，配置域的计费方案（因为在RADIUS服务器中，认证与授权是绑定在一起的，所以无需另外配置授权方案）。

（2）在17.2.4节表17-5中第7步后要添加 一条radius-server template-name命令指定域要使用的RADIUS服务器模板。这需要与在下面17.3.2节配置的RADIUS服务器模板一致，因为RADIUS服务器模板也只有在对应的域的AAA方案中绑定后才能得到应用。

17.3.1 配置AAA方案

在 RADIUS 方式中要配置采用 RADIUS 认证、授权和计费方式。配置认证模式为RADIUS认证时还可以配置本地认证或不认证为备份认证。配置备份认证（授权）可以避免单一认证（授权）模式无响应（不包括认证没通过的情况 ）而造成的认证失败。同理，配置计费模式为RADIUS计费时还可以配置不计费模式为备份计费。具体的配置步骤如表17-6所示。

表17-6 RADIUS认证、计费中的AAA方案配置步骤

步骤	命令	说明
1	**system-view** 例如：<HUAWEI> **system-view**	进入系统视图
2	**aaa** 例如：[HUAWEI] **aaa**	进入 AAA 视图
配置 AAA 认证方案		
3	**authentication-scheme** *authentication-scheme-name* 例如：[HUAWEI-aaa] **authentication-scheme** scheme0	创建一个认证方案，并进入认证方案视图或直接进入一个已存在的认证方案视图。其他参见 17.2.1 节表 17-2 中的第 3 步说明

（续表）

步骤	命令	说明
4	**authentication-mode radius** [**none**] 例如：[HUAWEI-aaa-authen-scheme0]**authentication-mode radius**	配置认证模式为 RADIUS 认证，同时选择可选项"**none**"时，表示不进行认证，也可理解为直接让用户通过认证。如果想要同时配置本地认证方式为备份认证方式，则可配置 **authentication-mode radius local** 命令 【说明】如果在一个认证方案中使用多种认证模式，则认证模式的执行顺序为配置的先后顺序。**只有在当前认证模式没有响应的情况下，才会采用下一种认证模式；如果在当前认证模式认证失败，则不会跳转到下一个认证方案进行认证** 缺省情况下，认证模式本地认证，可用 **undo authentication-mode** 命令恢复当前认证方案使用的认证模式为缺省的本地认证模式
5	**authentication-super** { **hwtacacs** | **radius** | **super** } * [**none**] 例如：[HUAWEI-aaa-authen-scheme0]**authentication-super radius**	（可选）配置使用当前认证方案的用户级别提升时的认证模式。这与本地认证方式中的用户级别提升时的认证方法的配置方法是一样，其他说明参见表 17-2 中的第 5 步
6	**quit** 例如：[HUAWEI-aaa-authen-scheme0] **quit**	退出认证方案视图，返回 AAA 视图
7	**domainname-parse-direction** { **left-to-right** | **right-to-left** } 例如：[HUAWEI-aaa] **domainname-parse-direction left-to-right**	（可选）配置用户名和域名解析的方向。其他说明参见表 17-2 中的第 7 步
配置 AAA 计费方案		
8	**accounting-scheme** *accounting-scheme-name* 例如：[HUAWEI-aaa] **accounting-scheme** scheme1	创建一个计费方案，并进入计费方案视图或直接进入一个已存在的计费方案视图。其他参见 17.2.1 节表 17-2 中的第 3 步说明，只不过这里创建的是计费方案 缺省情况下，设备中有一个计费方案，计费方案名称是 **default**，不能删除，只能修改，可用 **undo authorization-scheme** *accounting-scheme-name* 命令删除指方案定的计费
9	**accounting-mode radius** 例如：[HUAWEI-aaa-accounting-scheme1] **accounting-mode radius**	配置计费模式为 **radius**。用户上线时，经过认证和授权则开始计费；用户下线时结束计费。担当 AAA 客户端的接入设备将计费报文上送给计费服务器，其中计费报文中记录了用户在线的时间 缺省情况下，计费模式采用不计费模式 **none**（不计费），可用 **undo accounting-mode** 命令恢复当前计费方案使用缺省的不计费模式
10	**accounting start-fail** { **online** | **offline** } 例如：[HUAWEI-aaa-accounting-scheme1]**accounting start-fail online**	（可选）配置开始计费失败策略。命令中的选项说明如下。 （1）**online**：二选一选项，指定开始计费失败策略为如果开始计费失败，仍允许用户上线 （2）**offline**：二选一选项，指定开始计费失败策略为如果开始计费失败，拒绝用户上线 缺省情况下，如果初始计费失败，不允许用户上线，可用 **undo accounting start-fail** 命令恢复计费失败策略为缺省情况

（续表）

步骤	命令	说明
11	**accounting realtime** *interval* 例如：[HUAWEI-aaa-accounting-scheme1]**accounting realtime** 60	（可选）使能实时计费功能，并设置实时计费时间间隔。参数 *interval* 用来指定实时计费的时间间隔，取值范围为 0～65535 的整数分钟 配置实时计费后，设备向计费服务器**定时**发送实时计费报文，计费服务器收到实时计费报文后才进行计费。如果设备检测到付费用户下线，则停止发送实时计费报文，计费服务器终止计费，从而减小了计费误差 缺省情况下，设备按时长计费，未使能实时计费功能，没有设置实时计费间隔，可用 **undo accounting realtime** 命令去使能实时计费功能
12	**accounting interim-fail** [**max-times** *times*] { **online** | **offline** } 例如：[HUAWEI-aaa-accounting-scheme1] **accounting interim-fail max-times** 5 **online**	（可选）配置允许设备发送的实时计费请求最大无响应次数，以及实时计费失败后采取的策略。命令中的参数和选项说明如下。 （1）**max-times** *times*：可选参数，指定允许实时计费请求最大无响应次数，取值范围为 1～255 的整数。当实时计费请求最大无响应次数达到此最大值时，如果下一次计费请求仍然没有响应，设备认为计费失败，对付费用户采用实时计费失败策略 （2）**online**：二选一选项，指定实时计费失败后采取的策略为 online，即如果实时计费失败，仍允许用户上线 （3）**offline**：二选一选项，指定实时计费失败后采取的策略为 offline，即如果实时计费失败，拒绝用户上线 缺省情况下，允许的实时计费请求最大无响应次数为 3 次，实时计费失败后仍保持付费用户在线，可用 **undo accounting interim-fail** 命令恢复缺省配置

17.3.2 配置RADIUS服务器模板

配置RADIUS服务器模板关键是用来配置与RADIUS服务器进行通信的相关参数，如RADIUS服务器的IP地址和端口号，与RADIUS服务器通信时所使用的共享密钥等。像RADIUS用户名格式、流量计算单位、RADIUS请求报文的超时重传次数等参数都有缺省配置，用户可以根据实际需要进行修改。RADIUS服务器模板也是要在对应的域的AAA方案中绑定才能得到应用。

RADIUS 服务器模板下的配置如 RADIUS 用户名格式、RADIUS 共享密钥等要与RADIUS服务器上的对应配置一致。RADIUS服务器有许多种方案（各种服务器操作系统都提供这一功能），具体参见相关文档。RADIUS服务器模板的具体配置步骤如表17-7所示。

表17-7 RADIUS服务器模板的配置步骤

<table>
<tr><th>步骤</th><th>命令</th><th>说明</th></tr>
<tr><td>1</td><td>**system-view**
例如：<HUAWEI> **system-view**</td><td>进入系统视图</td></tr>
<tr><td>2</td><td>**radius-server authorization** *ip-address* [**vpn-instance** *vpn-instance-name*] { **server-group** *group-name* | **shared-key** {**cipher** | **simple** } *key-string* }* [**ack-reserved-interval** *interval*]
例如：[HUAWEI] **radius-server authorization** 100.1.1.116 **server-group** template1 **shared-key cipher** hello</td><td>（可选）配置 RADIUS 服务器模板中的 RADIUS 授权服务器，包括服务器的 IP 地址和共享密钥。命令中的参数和选项说明如下。
（1）*ip-address*：指定 RADIUS 授权服务器的 IP 地址。因为在 RADIUS 中，认证和授权功能是同时启用的，所**以这里的授权服务器 IP 地址与本表第 4 步的 RADIUS 主用认证服务器的 IP 地址必须一致**
（2）*vpn-instance-name*：可选参数，指定要绑定的 VPN 实例名称，为 1～31 个字符，以英文字母 a～z 或 A～Z 开始，可以是英文字母、数字、连字符“-”或下划线的组合，区分大小写
（3）*group-name*：可多选参数，指定 RADIUS 授权服务器对应的 RADIUS 服务器模板名称，为 1～32 个字符，不支持空格，区分大小写
（4）**cipher**：二选一选项，指定以密文形式显示用户口令
（5）**simple**：二选一选项，指定以明文形式显示用户口令
（6）*key-string*：可多选参数，指定与 RADIUS 授权服务器通信的共享密钥，如果选择 **simple** 选项，则必须是明文密码，为 1～16 位字符串；如果选择 **cipher** 选项，则既可以是 32 位的密文密码，也可以是 1～16 位的明文密码。不支持空格、单引号和问号，区分大小写。**通常与本表第 8 步配置的密钥是一致的**
（7）*interval*：可选参数，指定授权响应报文的保留时长，取值范围为 0～300 的整数秒。缺省值是 0 秒，不保留。如果要保留 RADIUS 授权响应报文以用于响应 RADIUS 授权服务器的重传报文，需要配置授权回应报文保留时长
缺省情况下，没有配置 RADIUS 授权服务器，可用 **undo radius-server authorization** *ip-address* [**vpn-instance** *vpn-instance-name*] 命令删除 RADIUS 服务器模板中的 RADIUS 授权服务器的相关配置</td></tr>
<tr><td>3</td><td>**radius-server template** *template-name*
例如：[HUAWEI] **radius-server template** template1</td><td>创建 RADIUS 服务器模板，并进入 RADIUS 服务器模板视图。参数 *template-name* 用来指定要进入的 RADIUS 服务器模板名称。不同系列允许创建的 RADIUS 服务器模板数不一样，具体参见相应产品文档说明
缺省情况下，设备上没有 RADIUS 服务器模板，可用 **undo radius-server template** *template-name* 命令删除指定的 RADIUS 服务器模板</td></tr>
<tr><td>4</td><td>**radius-server authentication** *ip-address port* [**vpn-instance** *vpn-instance-name* | **source** { **loopback** *interface-number* | **ip-address** *ip-address* }]*
例如：[HUAWEI-radius-template1] **radius-server authentication** 10.163.155.13 1813 **source loopback** 10</td><td>配置RADIUS主用认证服务器。命令中的参数说明如下。
（1）*ip-address port*：指定 RADIUS 认证服务器的 IP 地址和端口号，端口号的取值范围为 1～65 535 的整数
（2）*vpn-instance-name*：可多选参数，指定要绑定的 VPN 实例名称
（3）**loopback** *interface-number*：可多选参数，指定作为源接口的的 Loopback 接口编号，取值范围为 0～1 023 的整数</td></tr>
</table>

（续表）

<table>
<tr><th>步骤</th><th>命令</th><th>说明</th></tr>
<tr><td>4</td><td>radius-server authentication ip-address port [vpn-instance vpn-instance-name | source { loopback interface-number | ip-address ip-address }] *
例如：[HUAWEI-radius-template1] radius-server authentication 10.163.155.13 1813 source loopback 10</td><td>• ip-address ip-address：可多选参数，指定作为向 RADIUS 认证服务器发送 RADIUS 报文时使用的源 IP 地址。如果没有配置此参数，则使用前面指定的 Loopback 接口的 IP 地址作为向 RADIUS 认证服务器发送 RADIUS 报文时使用的源 IP 地址
缺省情况下，RADIUS 主用认证服务器的 IP 地址为 0.0.0.0，端口号为 0，可用 undo radius-server authentication [ip-address port [vpn-instance vpn- instance-name]] [source { loopback interface-number | ip-address ip-address }] 命令删除指定的 RADIUS 主用认证服务器配置</td></tr>
<tr><td>5</td><td>radius-server authentication ip-address port [vpn-instance vpn-instance-name | source { loopback interface-number |ip-address ip-address }] *secondary
例如：[HUAWEI-radius-template1] radius-server authentication 10.163.155.15 1813 source loopback 10 secondary</td><td>（可选）配置 RADIUS 备用认证服务器。其他说明参见上一步主 RADIUS 认证服务器配置
缺省情况下，RADIUS 主用认证服务器的 IP 地址为 0.0.0.0，端口号为 0，可用 undo radius-server authentication [ip-address port [vpn-instance vpn-instance-name]] [source { loopback interface-number | ip-address ip-address }] secondary 命令删除指定的RADIUS 备用认证服务器配置</td></tr>
<tr><td>6</td><td>radius-server accounting ip-address port [vpn-instance vpn-instance-name | source { loopback interface-number | ip-addressip-address }] *
例如：HUAWEI-radius-template1] radius-server accounting 10.163.155.13 1812 source loopback 10</td><td>配置 RADIUS 主用计费服务器。命令中的参数与前面第 4 步中 RADIUS 主用认证服务器中的对应参数一样，只不过这里指定的 RADIUS 主计费服务器中的对应参数值，参见即可
通常 RADIUS 主用计费服务器与 RADIUS 主用认证服务器在同一台主机上，所以两者的 IP 地址通常是一样的，端口也可以一样
缺省情况下，RADIUS 主用计费服务器的 IP 地址为 0.0.0.0，端口号为 0，可用 undo radius-server accounting [ip-address port [vpn-instance vpn-instance-name]] [source { loopback interface-number | ip-address ip-address }] 命令删除指定的 RADIUS 主用计费服务器的相关配置</td></tr>
<tr><td>7</td><td>radius-server accounting ip-address port [vpn-instance vpn-instance-name | source { loopback interface-number | ip-address ip-address }] * secondary
例如：[HUAWEI-radius-template1] radius-server accounting 10.163.155.15 1812 source loopback 10 secondary</td><td>（可选）配置 RADIUS 备用计费服务器。参数同样可参见前面第 4 步中的 RADIUS 主用认证服务器中的对应参数
通常 RADIUS 备用计费服务器与 RADIUS 备用认证服务器是在同一台主机上，所以两者的 IP 地址通常是一样的，端口也可以一样
缺省情况下，RADIUS 备份计费服务器的 IP 地址为 0.0.0.0，端口号为 0，可用 undo radius-server accounting [ip-address port [vpn-instance vpn-instance-name]] [source { loopback interface-number | ip-address ip-address }] secondary 命令删除指定的 RADIUS 备用计费服务器的相关配置</td></tr>
<tr><td>8</td><td>radius-server shared-key [cipher | simple] key-string
例如：[HUAWEI-radius-template1] radius-server shared-key cipher huawei</td><td>（可选)配置担当 AAA 客户端的本地接入设备与 RADIUS 服务器（包括认证和计费服务器）通信的共享密钥。命令中的参数可参见本表第 2 步的中对应参数说明
【说明】设备和 RADIUS 服务器在发送认证报文时，对口令等重要信息使用 MD5 加密，确保认证信息在网络中传输的安全性。为了确保认证双方身份的合法性，要求设备上配置的密钥与 RADIUS 认证服务器的密钥相同。如果配置密钥时不带 simple 或 cipher 关键字，则按密文形式处理
缺省情况下，RADIUS 共享密钥是 huawei，采用明文形式显示用户口令，可用 undo radius-server shared-key 命令恢复缺省配置</td></tr>
</table>

（续表）

<table>
<tr><th>步骤</th><th>命令</th><th>说明</th></tr>
<tr><td>9</td><td>radius-server user-name domain-included
例如：[HUAWEI-radius-template1] radius-server user-name domain-included</td><td>（可选）配置设备向 RADIUS 服务器发送的报文中的用户名包含域名。用户名通常采用“纯用户名@域名”格式，@后面的部分为域名。这里@表示域名分隔符，域名分隔符也可以是 \/:<>|'% 中的一个。但只有当该 RADIUS 服务器模板没有用户使用时，才能改变此配置
缺省情况下，RADIUS 用户名中包含域名，即设备会把用户名和域名及域名分隔符一起发送给 RADIUS 服务器进行认证。如果 RADIUS 服务器不接受带域名的用户名，可以执行命令 undo radius-server user-name domain-included，设备会将用户名中的域名去掉，再发送给 RADIUS 服务器</td></tr>
<tr><td>10</td><td>radius-server traffic-unit { byte | kbyte | mbyte | gbyte }
例如：[HUAWEI-radius-template1] radius-server traffic-unit kbyte</td><td>（可选）配置 RADIUS 计费服务器计费时所采用的流量统计单位。命令中的选项说明如下。
（1）byte：多选一选项，指定以字节为流量单位
（2）kbyte：多选一选项，指定以千字节为流量单位
（3）mbyte：多选一选项，指定以兆字节为流量单位
（4）gbyte：多选一选项，指定以吉字节为流量单位
由于不同的 RADIUS 服务器使用的流量统计单位可能不同，因此，需要在设备上针对每一个 RADIUS 服务器组设置流量单位，和 RADIUS 服务器保持一致。但只有当该 RADIUS 服务器模板没有用户使用时，才能改变流量单位的配置
缺省情况下，设备以字节（byte）作为 RADIUS 流量单位，可用 undo radius-server traffic-unit 命令恢复为缺省配置</td></tr>
<tr><td>11</td><td>radius-server { retransmit retry-times | timeout time-value }*
例如：[HUAWEI-radius-template1] radius-server retransmit 4 timeout 8</td><td>（可选）配置 RADIUS 请求报文允许的超时重传次数和超时时间。命令中的参数说明如下。
（1）retry-times：可多选参数，指定允许的 RADIUS 请求报文超时重传次数，取值范围为 1～5 的整数
（2）time-value：可多选参数，指定 RADIUS 请求报文的超时时间，取值范围为 3～10 的整数秒，缺省值是 5 秒
缺省情况下，RADIUS 请求报文的超时重传次数为 3，超时时间是 5s</td></tr>
<tr><td>12</td><td>radius-server nas-port-format { new | old }
例如：[HUAWEI-radius-template1] radius-server nas-port-format new</td><td>（可选）配置 RADIUS 服务器的 NAS 端口形式。命令中的选项说明如下。
（1）new：二选一选项，指定采用新的 NAS 端口形式
（2）old：二选一选项，指定采用旧的 NAS 端口形式
【说明】NAS 端口两种形式的区别主要在于以太接入用户的物理端口：选择 new 时，端口形式为槽位号（8 位）＋子槽位号（4 位）＋端口号（8 位）＋VLAN ID（12 位）；选择 old 时，端口形式为槽位号（12 位）＋端口号（8 位）＋VLAN ID（12 位）。但对于 ADSL 接入用户，不受此命令的影响，端口形式都是槽位号（4 位）＋子槽位号（2 位）＋端口号（2 位）＋VPI（8 位）＋VCI（16 位）。NAS 端口和 NAS 端口 ID 形式都属于华为公司内部扩展的属性，仅用于华为公司设备之间的互通和业务配合
缺省情况下，采用新的 NAS 端口形式，可用 undo radius-server nas-port-format 命令恢复缺省的 NAS 端口形式</td></tr>
</table>

（续表）

步骤	命令	说明
13	**radius-server nas-port-id-format** { **new** | **old** } 例如：[HUAWEI-radius-template1] **radius-server nas-port-id-format new**	（可选）配置 RADIUS 服务器的 NAS 端口 ID 形式。命令中的选项说明如下。 （1）**new**：二选一选项，指定采用新的 NAS 端口 ID 形式 （2）**old**：二选一选项，指定采用旧的 NAS 端口 ID 形式 【说明】NAS 端口 ID 两种形式的说明如下。 （1）选择 **new** 时，对于以太接入用户，NAS 端口 ID 形式为：slot=XX；subslot=XX；port=XXX；VLANID=XXXX；，其中，Slot 取值范围为 0~15 的整数，Subslot 取值范围为 0~15 的整数，Port 取值范围为 0~255 的整数，VLANID 取值范围为 1~4 094 的整数；对于 ADSL 接入用户，NAS 端口 ID 形式为：slot=XX；subslot=X；port=X；VPI=XXX；VCI=XXXXX；其中，Slot 取值范围为 0~15 的整数，Subslot 取值范围为 0~9 的整数，Port 取值范围为 0~9 的整数，VPI 取值范围为 0~255 的整数，VCI 取值范围为 0~65 535 的整数 （2）选择 **old** 时，对于以太接入用户，NAS 端口 ID 形式为：端口号（两个字符）+子槽号（两个字节）+卡号（三个字节）+VLANID（9 个字符）；对于 ADSL 接入用户，NAS 端口 ID 形式为：端口号（两个字符）+子槽号（两个字节）+卡号（三个字节）+VPI（8 个字符）+VCI（16 个字符），字节数不够的，在前面补零 **NAS 端口和 NAS 端口 ID 形式都属于华为公司内部扩展的属性，仅用于华为公司设备之间的互通和业务配合** 缺省情况下，采用新的 NAS 端口 ID 形式，可用 **undo radius-server nas-port-id-format** 命令恢复为缺省的 NAS 端口 ID 形式
14	**radius-attribute nas-ip** *ip-address* 例如：[HUAWEI-radius-template1] **radius-attribute nas-ip** 10.163.155.13	（可选）配置 NAS 发送 RADIUS 报文使用的 NAS-IP-Address 属性。参数 *ip-address* 用来指定设备发送 RADIUS 报文使用的 NAS-IP-Address 属性值，也就是发送 RADIUS 报文的源 IP 地址 缺省情况下，使用 NAS 源 IP 地址（也就是在本表第 4 步指定的 loopback 接口 IP 地址）作为 NAS-IP-Address 属性的值，可用 **undo radius-attribute nas-ip** 命令删除配置的 NAS-IP-Address 属性
15	**radius-server accounting-stop-packet resend** [*resend-times*] 例如：[HUAWEI-radius-template1] **radius-server accounting-stop-packet resend** 3	（可选）使能 RADIUS 计费结束报文的重传功能，并配置可重发的计费停止报文个数。可选参数 *resend-times* 用来指定可重发的计费停止报文个数，取值范围为 1~300 的整数。若执行本命令时不输入此可选参数，则缺省为 100 缺省情况下，计费停止报文的重发次数为 0，即计费停止报文不重发，可用 **undo radius-server accounting-stop-packet** 命令恢复为缺省情况
16	**radius-server dead-time** *dead-time* 例如：[HUAWEI-radius-template1] **radius-server dead-time** 1	（可选）配置 RADIUS 主用服务器恢复激活状态的时间，取值范围为 1~65 535 的整数分钟 【说明】当设备将 RADIUS 服务器的状态置为 Down 后，等待本命令配置的时间后，设备会重新将 RADIUS 服务器的状态置为 UP，并尝试和 RADIUS 服务器重新建立连接。如果连接失败，设备重新将 RADIUS 服务器的状态置为 Down。**但只有当该 RADIUS 模板没有用户使用时，才能改变此配置** 缺省情况下，主用服务器恢复激活状态的时间为 5min，可用 **undo radius-server dead-time** 命令将主用服务器恢复激活状态的时间恢复为缺省值

（续表）

步骤	命令	说明
17	**return** 例如：[HUAWEI-radius-template1] **return**	返回用户视图
18	**test-aaa** *user-name user-password* **radius-template** *template-name* [**chap** | **pap**] 例如：[HUAWEI-radius-template1] **test-aaa** user1@mydomain userkey **radius-template** template1 **pap**	（可选）测试用户是否能够通过 RADIUS 认证。命令中的参数和选项说明如下： （1）*user-name user-password*：指定要测试的用户名和密码 （2）*template-name*：指定测试的 RADIUS 服务器模板名称 （3）**chap**：二选一选项，指定认证方式为 CHAP 认证 （4）**pap**：二选一选项，指定认证方式为 PAP 认证 如果某个用户无法通过认证，可以在设备上执行本命令定位故障：如果测试结果表明该用户可以通过 RADIUS 认证，则证明故障出现在接入认证；如果测试结果表明该用户无法通过 RADIUS 认证，则证明故障出现在 RADIUS 认证

17.3.3 RADIUS认证、授权和计费配置示例

本示例拓扑结构如图17-4所示，用户同处于huawei域，通过SwitchA访问网络。SwitchB 作为目的网络

（Destination Network）的接入服务器（NAS）。现要在 SwitchB上采用以下RADIUS方式的AAA方案控制用户访问目的网络。

（1）为了提高认证的可靠性，SwitchB对接入用户先用RADIUS服务器进行认证，如果认证没有响应，再使用本地认证。

（2）RADIUS 服务器 129.7.66.66/24 作为主用认证服务器和计费服务器，RADIUS服务器129.7.66.67/24作为备用认证服务器和计费服务器，认证端口号缺省为1812，计费端口号缺省为1813。

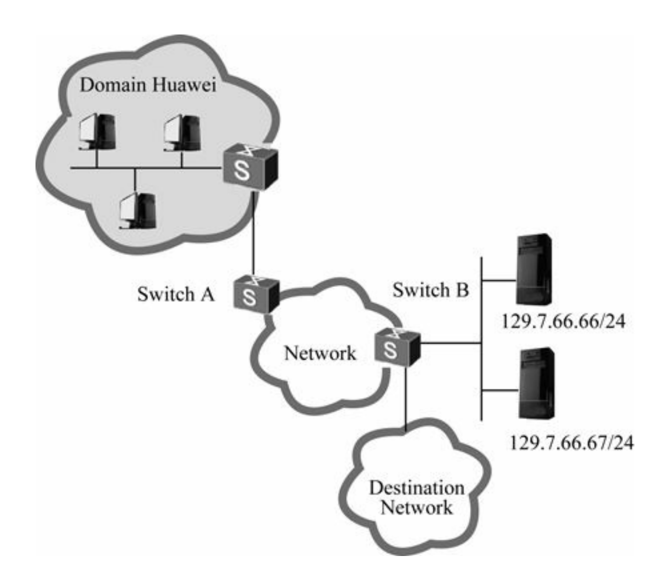


图17-4 RADIUS方式认证、授权和计费配置示例拓扑结构

1．基本配置思路

根据图17-3（b）所示的流程，再结合本示例的具体要求得出以下基本配置思路（均在SwitchB上配置，没有特别的业务方案要求，所以可无需配置业务方案，直接采用缺 省配置即可）：

（1）配置AAA方案，包括RADIUS认证方案和计费方案。

（2）配置RADIUS服务器模板。

（3）在huawei域下绑定上面的RADIUS认证、计费方案和RADIUS服务器模板。

2．具体配置步骤

（1）配置RADIUS认证方案和计费方案。根据示例的要求，以本地认证作为备份认证方式。

[HUAWEI] aaa

[HUAWEI-aaa] authentication-scheme auth #---配置认证方案名为auth

[HUAWEI-aaa-authen-auth] authentication-mode radius local #---配置认证模式为先进行RADIUS认证，RADIUS认证服务器无响应后再进行本地认证

[HUAWEI-aaa-authen-auth] quit

[HUAWEI-aaa] accounting-scheme abc #---配置计费方案abc

[HUAWEI-aaa-accounting-abc] accounting-mode radius #---配置计费模式为RADIUS计费模式

[HUAWEI-aaa-accounting-abc] accounting start-fail online #---配置当开始计费失败时，允许用户上线

[HUAWEI-aaa-accounting-abc] quit

（2）配置RADIUS服务器模板。

<HUAWEI>system-view

[HUAWEI] radius-server template shiva   #---配置RADIUS服务器模板 shiva

[HUAWEI-radius-shiva] radius-server authentication 129.7.66.66 1812   #---配置RADIUS主用认证服务器的IP地址和端口

[HUAWEI-radius-shiva] radius-server accounting129.7.66.66 1813   #---配置RADIUS主用计费服务器的 IP地址和端口

[HUAWEI-radius-shiva] radius-server authentication 129.7.66.67 1812 secondary   #---配置RADIUS备用认证服务器的IP地址和端口

[HUAWEI-radius-shiva] radius-server accounting129.7.66.67 1813 secondary   #---配置RADIUS备用计费服务器的 IP地址和端口

[HUAWEI-radius-shiva] radius-server shared-key cipherhello   #---配置RADIUS服务器的共享密钥为hello

[HUAWEI-radius-shiva] radius-server retransmit 2   #---配置设备向RADIUS服务器发送请求报文的超时重传次数为2

[HUAWEI-radius-shiva] quit

（3）配置huawei 域，并在域下绑定以上配置的认证方案、计费方案和RADIUS服务器模板。

[HUAWEI-aaa] domain huawei

[HUAWEI-aaa-domain-huawei] authentication-scheme auth

[HUAWEI-aaa-domain-huawei] accounting-scheme abc

[HUAWEI-aaa-domain-huawei] radius-server shiva

配置好后在SwitchB上执行display radius-server configuration template命令可以查看到该RADIUS服务器模板的配置与上述配置是一致的。

17.4 HWTACACS方式认证、授权和计费配置

HWTACACS协议与RADIUS协议类似，主要是通过C/S模式与HWTACACS服务器通信来实现对接入用户进行认证、授权和计费。但与RADIUS相比，HWTACACS具有更加可靠的传输和加密特性，更加适合于安全控制。采用HWTACACS方式进行认证、授权、计费可以防止非法用户对网络的攻击，HWTACACS 还支持对命令行进行授权，比RADIUS更适用于进行安全控制。

本节要介绍使用HWTACACS协议对接入用户进行认证、授权和计费的配置方法（不包括**HWTACACS**服务器中用户账户信息和用户授权属性等自身的配置 ）。

说明

根据图 17-3（c）可以得知 HWTACACS 认证、授权和计费方式的配置流程。在HWTACACS 方式的配置中，与本地方式的配置中也主要不同在于前面两项配置任务，本节仅介绍这两项配置任务的具体配置方法和步骤。

第三项的“配置业务方案”和与前面17.2.3节介绍的本地方式中的“配置业务方案”的配置方法和步骤完全一样，不同的只是这里配置的业务方案适用于HWTACACS服务器认证、授权和计费。

第四项的“配置域的AAA方案”和17.2.4节介绍的本地方式中的“配置域的AAA方案”的配置方法与步骤总体也一样，两个的不同之处如下。

（1）在 17.2.4 节表 17-5 中第 5 步之后要添加 一条 accounting-scheme accounting-scheme-name命令来配置域的计费方案（在HWTACACS服务器中，认证、授权和计费功能都是分开的，所以需要分别为它们配置相应使用的业务方案）。

（2）在 17.2.4节表 17-5中第 7步后要添加 一条hwtacacs-server template-name命令指定域要使用的HWTACACS服务器模板。这需要与在下面17.4.2节配置的HWTACACS服务器模板一致。

17.4.1 配置AAA方案

采用HWTACACS方式时，需要配置HWTACACS认证、授权和计费模式，同样还可以配置本地认证、授权或不认证、不授权为备份认证或授权模式。配置备份认证可以避免单一认证或授权模式无响应而造成的认证或授权失败。HWTACACS方式AAA方案的配置步骤如表17-8所示。

表17-8 HWTACACS方式AAA方案的配置步骤

步骤	命令	说明
1	**system-view** 例如：<HUAWEI> **system-view**	进入系统视图
2	**aaa** 例如：[HUAWEI] **aaa**	进入 AAA 视图
配置 AAA 认证方案		
3	**authentication-scheme** *authentication-scheme-name* 例如：[HUAWEI-aaa] **authentication-scheme** scheme0	创建一个认证方案，并进入认证方案视图或直接进入一个已存在的认证方案视图。其他参见表 17-2 中的第 3 步说明
4	**authentication-mode hwtacacs** [**none**] 例如：[HUAWEI-aaa-authen-scheme0]**authentication-mode hwtacacs**	配置认证模式为 **hwtacacs** 认证。选择可选项“**none**”时表示不进行认证，也可理解为直接让用户通过认证。如果想要同时配置本地认证方式为备份认证方式，则可配置 **authentication-mode hwtacacs local** 命令 【说明】如果在一个认证方案中使用多种认证模式，则认证模式的执行顺序为配置的先后顺序。只有在当前认证模式没有响应的情况下，才会采用下一种认证模式；如果在当前认证模式认证失败，则不会跳转到下一个认证方案进行认证 缺省情况下，认证模式就为本地认证，可用 **undo authentication-mode** 命令恢复当前认证方案使用的认证模式为缺省的本地认证模式

（续表）

步骤	命令	说明
5	**authentication-super** { **hwtacacs** | **radius** | **super** }* [**none**] 例如：[HUAWEI-aaa-authen-scheme0]**authentication-super radius**	（可选）配置使用当前认证方案的用户级别提升时的认证模式。这与本地认证方式中的用户级别提升时的认证方法的配置方法一样，其他说明参见表 17-2 中的第 5 步
6	**quit** 例如：[HUAWEI-aaa-authen-scheme0] **quit**	退出认证方案视图，返回 AAA 视图
7	**domainname-parse-direction** { **left-to-right** | **right-to-left** } 例如：[HUAWEI-aaa] **domainname-parse-direction left-to-right**	（可选）配置用户名和域名解析的方向。其他说明参见表 17-2 中的第 7 步
8	**quit** 例如：[HUAWEI-aaa] **quit**	退出 AAA 视图，返回系统视图
9	**aaa-authen-bypass enable time** *time-value* 例如：[HUAWEI] **aaa-authen-bypass enable time** 2	配置认证旁路时间，取值范围为 1～1 440 整数分钟 **【说明】**使能此功能后，如果远端认证无响应，则在配置的旁路时间内直接跳过无响应的远端认证，直接跳转到第 4 步所配置的下一个认证方式，如果未配置下一个认证方式，则按失败处理 缺省情况下，未配置认证旁路时间，可用 **undo aaa-authen-bypass enable** 命令取消配置认证旁路时间
	配置 AAA 授权方案	
10	**aaa** 例如：[HUAWEI] **aaa**	进入 AAA 视图
11	**authorization-scheme** *authorization-scheme-name* 例如：[HUAWEI-aaa] **authorization-scheme** scheme1	创建一个授权方案，并进入授权方案视图或直接进入一个已存在的授权方案视图。其他说明参见表 17-2 中的第 8 步
12	**authorization-mode**{ **hwtacacs** | **local** }* [**none**] 例如：[HUAWEI-aaa-author-scheme1] **authorization-mode local**	配置授权模式。命令中的选项说明如下。 （1）**hwtacacs**：可多选项，指定授权模式为 HWTACACS 授权模式。如果采用 HWTACACS 授权模式，必须配置 HWTACACS 服务器模板，然后在用户所属域的视图下应用该服务器模板 （2）**local**：可多选项，指定授权模式为本地授权模式。如果同时选择了 **hwtacacs** 选项，则授权模式的执行顺序为配置的先后顺序。只有在当前授权模式没有响应的情况下，才会采用下一种授权模式；如果当前授权模式失败，则不会采用下一种授权模式进行授权 （3）**none**：可选项，指定授权模式为无需授权，直接为用户授权 缺省情况下，授权模式为本地授权模式，可用 **undo authorization-mode** 命令恢复当前授权方案使用的授权模式为缺省的本地授权模式
13	**authorization-cmd** *privilege-level* **hwtacacs** [**local**] [**none**] 例如：[HUAWEI-aaa-author-scheme1] **authorization-cmd** 2 **hwtacacs**	（可选）为指定级别的用户配置按命令行授权（也就是使用 **command-privilege level** *level* **view** *view-name command- key* 命令将指定的命令行指定为对应的命令级别，**具体参见本书第 2 章 2.1.4 节**）。命令中的参数和选项说明如下。 （1）*privilege-level*：指定要进行命令行授权的用户级别，取值范围为 0～15 的整数 （2）**hwtacacs**：指定按命令行授权模式为 HWTACACS 模式。如果使能按 HWTACACS 模式进行命令行授权，必须配置 HWTACACS 服务器模板，然后在用户所属域的视图下应用该服务器模板

（续表）

步骤	命令	说明
13	**authorization-cmd** *privilege-level* **hwtacacs** [**local**] [**none**] 例如：[HUAWEI-aaa-author-scheme1] **authorization-cmd** 2 **hwtacacs**	（3）**local**：可选项，指定当 HWTACACS 服务器出现故障导致授权失败，将授权方式转为本地授权 缺省情况下，0～15 级用户都没有配置按命令行授权，直接使用该用户级别可对应的命令级别，可用 **undo authorization-mode** 命令恢复当前授权方案使用的授权模式为缺省配置。但使能按命令行授权功能的授权方案被域引用后，如果执行 **undo authorization-cmd** 命令，将导致该域相应级别的在线用户无法执行任何命令（**quit** 命令除外）。此时用户需要重新登录
14	**quit** 例如：[HUAWEI-aaa-author-scheme1] **quit**	退出授权方案视图，返回 AAA 视图
15	**quit** 例如：[HUAWEI-aaa] **quit**	退出 AAA 视图，返回系统视图
16	**aaa-author-bypass enable time** *time-value* 例如：[HUAWEI] **aaa-author-bypass enable time** 2	（可选）配置授权旁路时间，取值范围为 1～1 440 整数分钟 【说明】使能此功能后，如果远端授权无响应，则在配置的旁路时间内直接跳过无响应的远端授权，直接跳转到第 12 步所配置的下一个授权模式，如果未配置下一个授权模式，则按失败处理 缺省情况下，未配置授权旁路时间，可用 **undo aaa-author-bypass enable** 命令取消配置授权旁路时间
17	**aaa-author-cmd-bypass enable time** *time-value* 例如：[HUAWEI] **aaa-author-cmd-bypass enable time** 2	（可选）配置命令行授权旁路时间，取值范围为 1～1 440 整数分钟 【说明】使能此功能后，如果远端命令行授权无响应，则在配置的旁路时间内直接跳过无响应的远端命令行授权，直接跳转到第 13 步所配置的下一个命令行授权模式，如果未配置下一个命令行授权模式，则按失败处理 缺省情况下，未配置命令行授权旁路时间，可用 **undo aaa-author-cmd-bypass enable** 命令取消配置命令行授权旁路时间
配置 AAA 计费方案		
18	**aaa** 例如：[HUAWEI] **aaa**	进入 AAA 视图
19	**accounting-scheme** *accounting-scheme-name* 例如：[HUAWEI-aaa] **accounting-scheme** scheme2	创建一个计费方案，并进入计费方案视图或直接进入一个已存在的计费方案视图。其他说明参见表 17-6 中的第 8 步
20	**accounting-mode hwtacacs** 例如：[HUAWEI-aaa-accounting-scheme2] **accounting-mode hwtacacs**	配置计费模式为 **hwtacacs**。用户上线时，经过认证和授权，计费开始；用户下线时，计费结束。担当 AAA 客户端的接入设备将计费报文上送给计费服务器，其中计费报文中记录了用户在线的时间 缺省情况下，计费模式采用不计费模式 **none**，可用 **undo accounting-mode** 命令恢复当前计费方案使用的计费模式为缺省的不计费模式
21	**accounting start-fail** { **online** | **offline** } 例如：[HUAWEI-aaa-accounting-scheme2]**accounting start-fail online**	（可选）配置开始计费失败策略。其他说明参见表 17-6 中的第 10 步

（续表）

步骤	命令	说明
22	**accounting realtime** *interval* 例如：[HUAWEI-aaa-accounting-scheme2]**accounting realtime** 60	（可选）使能实时计费，并设置实时计费时间间隔。其他说明参见表 17-6 中的第 11 步
23	**accounting interim-fail** [**max-times** *times*] { **online** | **offline** } 例如：[HUAWEI-aaa-accounting-scheme2] **accounting interim-fail max-times** 5 **online**	（可选）配置允许的实时计费请求最大无响应次数，以及实时计费失败后采取的策略。其他说明参见表 17-6 中的第 12 步

17.4.2 配置HWTACACS服务器模板

与17.3.2节介绍的RADIUS服务器模板配置一样，配置HWTACACS服务器模板中的关键步骤也是指定服务器的IP地址和端口号、HWTACACS共享密钥。其他的步骤如配置HWTACACS用户名格式、流量单位等都有缺省配置，用户可以根据实际需要进行修改。

HWTACACS服务器模板下配置的HWTACACS用户名格式、HWTACACS共享密钥等要与HWTACACS服务器上的对应配置一致。HWTACACS服务器模板的具体配置步骤如表17-9所示。

表17-9 HWTACACS服务器模板的配置步骤

步骤	命令	说明
1	**system-view** 例如：<HUAWEI> **system-view**	进入系统视图
2	**hwtacacs enable** 例如：[HUAWEI] **hwtacacs enable**	（可选）使能 HWTACACS 功能。但如果有用户正在进行 HWTACACS 认证或授权，或在线的用户使用 HWTACACS 计费，该命令执行不成功 缺省情况下，已使能 HWTACACS 功能，可用 **undo hwtacacs enable** 命令去使能 HWTACACS 功能
3	**hwtacacs-server template** *template-name* 例如：[HUAWEI] **hwtacacs-server template** template1	创建 HWTACACS 服务器模板，并进入 HWTACACS 服务器模板视图。参数 *template-name* 用来指定要进入的 HWTACACS 服务器模板名称，为 1～32 个字符，不支持空格，区分大小写，可以是英文字母、数字、连字符“-”或下划线的组合。不同系列允许创建的 HWTACACS 服务器模板数不一样，具体参见相应产品文档说明 缺省情况下，设备上没有 HWTACACS 服务器模板，可用 **undo hwtacacs-server template** *template-name* 命令删除指定的 HWTACACS 服务器模板
4	**hwtacacs-server authentication** *ip-address* [*port*] [**public-net** | **vpn-instance** *vpn-instance-name*] 例如：[HUAWEI-hwtacacs-template1]**hwtacacs-server authentication** 10.163.155.13 **vpn-instance** vpna	配置 HWTACACS 主用认证服务器。命令中的参数说明如下。 （1）*ip-address* [*port*]：指定 HWTACACS 认证服务器的 IP 地址和端口号，端口号的取值范围为 1～65 535 的整数 （2）**public-net**：二选一选项，指定在公网中连接 HWTACACS 认证服务器 （3）*vpn-instance-name*：二选一参数，指定要绑定的 VPN 实例名称 缺省情况下，HWTACACS 主用认证服务器的 IP 地址为 0.0.0.0，端口号为 0，不绑定 VPN 实例，可用 **undo hwtacacs-server authentication** *ip-address* [*port*] [**public-net** | **vpn-instance** *vpn-instance-name*] 命令删除指定的 HWTACACS 主用认证服务器配置

（续表）

步骤	命令	说明
5	**hwtacacs-server authentication** *ip-address* [*port*] [**public-net** | **vpn-instance** *vpn-instance-name*] **secondary** 例如：[HUAWEI-hwtacacs-template1] **hwtacacs-server authentication** 10.163.155.14 **vpn-instance** vpna **secondary**	（可选）配置 HWTACACS 备用认证服务器。其他说明参见上一步主 HWTACACS 认证服务器配置 缺省情况下，HWTACACS 备用认证服务器的 IP 地址为 0.0.0.0，端口号为 0，不绑定 VPN 实例，可用 **undo hwtacacs-server authentication** *ip-address* [*port*] [**public-net** | **vpn-instance** *vpn-instance-name*] **secondary** 命令删除指定的 HWTACACS 备用认证服务器配置
6	**hwtacacs-server authorization** *ip-address* [*port*] [**public-net** | **vpn-instance** *vpn-instance-name*] 例如：[HUAWEI-hwtacacs-template1]**hwtacacs-server authorization** 10.163.155.13 **vpn-instance** vpna	配置 HWTACACS 主用授权服务器。命令中的参数与前面第 4 步中 HWTACACS 主用认证服务器中的对应参数一样，只不过这里指定的 HWTACACS 主用授权服务器中的对应参数值，参见即可 **通常 HWTACACS 主用授权服务器与 HWTACACS 主用认证服务器是在同一台主机上，所以两者的 IP 地址通常是一样的，端口号也可以一样** 缺省情况下，HWTACACS 主用授权服务器的 IP 地址为 0.0.0.0，端口号为 0，不绑定 VPN 实例，可用 **undo hwtacacs-server authorization** *ip-address* [*port*] [**public-net** | **vpn-instance** *vpn-instance-name*] 命令删除指定的 HWTACACS 主用授权服务器的相关配置
7	**hwtacacs-server authorization** *ip-address* [*port*] [**public-net** | **vpn-instance** *vpn-instance-name*] **secondary** 例如：[HUAWEI-hwtacacs-template1]**hwtacacs-server authorization** 10.163.155.14 **vpn-instance** vpna **secondary**	（可选）配置 HWTACACS 备用授权服务器。命令中的参数与前面第 4 步中 HWTACACS 主用认证服务器中的对应参数一样，只不过这里指定的 HWTACACS 备用授权服务器中的对应参数值，参见即可 **通常 HWTACACS 备用授权服务器与 HWTACACS 备用认证服务器是在同一台主机上，所以两者的 IP 地址通常是一样的，端口号也可以一样** 缺省情况下，HWTACACS 备用授权服务器的 IP 地址为 0.0.0.0，端口号为 0，不绑定 VPN 实例，可用 **undo hwtacacs-server authorization** *ip-address* [*port*] [**public-net** | **vpn-instance** *vpn-instance-name*] **secondary** 命令删除指定的 HWTACACS 备用授权服务器的相关配置
8	**hwtacacs-server accounting** *ip-address* [*port*] [**public-net** | **vpn-instance** *vpn-instance-name*] 例如：[HUAWEI-hwtacacs-template1]**hwtacacs-server accounting** 10.163.155.13 49 **vpn-instance** vpna	配置 HWTACACS 主用计费服务器。参数同样可参见前面第 4 步中的 HWTACACS 主用认证服务器中的对应参数 **通常 HWTACACS 主用计费服务器与 HWTACACS 主用授权、计费服务器是在同一台主机上，所以三者的 IP 地址通常是一样的，端口号也可以一样** 缺省情况下，HWTACACS 主用计费服务器的 IP 地址为 0.0.0.0，端口号为 0，不绑定 VPN 实例，可用 **undo hwtacacs-server accounting** *ip-address* [*port*] [**public-net** | **vpn-instance** *vpn-instance-name*] 命令删除指定的 HWTACACS 主用计费服务器的相关配置
9	**hwtacacs-server accounting** *ip-address* [*port*] [**public-net** | **vpn-instance** *vpn-instance-name*] **secondary** 例如：[HUAWEI-hwtacacs-template1]**hwtacacs-server accounting** 10.163.155.14 49 **vpn-instance** vpna **secondary**	（可选）配置 HWTACACS 备用计费服务器。参数同样可参见前面第 4 步中的 HWTACACS 主用认证服务器中的对应参数 **通常 HWTACACS 备用计费服务器与 HWTACACS 备用授权、备用计费服务器是在同一台主机上，所以三者的 IP 地址通常是一样的，端口号也可以一样** 缺省情况下，HWTACACS 备用计费服务器的 IP 地址为 0.0.0.0，端口号为 0，不绑定 VPN 实例，可用 **undo hwtacacs-server accounting** *ip-address* [*port*] [**public-net** | **vpn-instance** *vpn-instance-name*] **secondary** 命令删除指定的 HWTACACS 备用计费服务器的相关配置

（续表）

步骤	命令	说明
10	**hwtacacs-server source-ip** *ip-address* 例如：[HUAWEI-hwtacacs-template1]**hwtacacs-server source-ip** 10.1.1.1	（可选）配置设备向 HWTACACS 服务器发送 HWTACACS 报文的源 IP 地址。指定 HWTACACS 源 IP 地址后，设备使用该 HWTACACS 服务器模板与服务器通信时，报文的源 IP 地址为指定的 IP 地址 缺省情况下，HWTACACS 的源 IP 地址是 0.0.0.0，此时设备使用实际出方向的接口的 IP 地址作为 HWTACACS 报文的源 IP 地址，可用 **undo hwtacacs-server source-ip** 命令将设备向 HWTACACS 服务器发送 HWTACACS 报文的源 IP 地址恢复为缺省值
11	**hwtacacs-server shared-key** [**cipher** | **simple**] *key-string* 例如：[HUAWEI-hwtacacs-template1] **hwtacacs-server shared-key cipher** hello	（可选）配置担当 AAA 客户端的本地接入设备与 HWTACACS 服务器通信的共享密钥。命令中的参数和选项说明如下。 （1）**cipher**：二选一选项，指定以密文形式显示用户口令 （2）**simple**：二选一选项，指定以明文形式显示用户口令 （3）*key-string*：可多选参数，指定与 RADIUS 认证服务器通信的共享密钥，如果选择 **simple** 选项，则必须是明文密码，为 1～255 位字符串；如果选择 **cipher** 选项，则既可以是 1～255 位明文密码，也可以为 20～392 位密文密码。如果配置密钥时不带 **simple** 或 **cipher** 关键字，则按密文形式处理 缺省情况下，没有配置 HWTACACS 服务器共享密钥，可用 **undo hwtacacs-server shared-key** 命令删除配置的 HWTACACS 服务器共享密钥
12	**hwtacacs-server user-name domain-included** 例如：[HUAWEI-hwtacacs-template1] **hwtacacs-server user-name domain-included**	（可选）配置设备向 HWTACACS 服务器发送的报文中的用户名包含域名。其他说明参见表 17-7 中的第 9 步 缺省情况下，HWTACACS 用户名中包含域名，即设备会把用户名和域名及域名分隔符一起发送给 HWTACACS 服务器进行认证。如果 HWTACACS 服务器不接受带域名的用户名，可以执行命令 **undo hwtacacs-server user-name domain-included**，设备会将用户名中的域名去掉，再发送给 HWTACACS 服务器
13	**hwtacacs-server traffic-unit** { **byte** | **kbyte** | **mbyte** | **gbyte** } 例如：[HUAWEI-hwtacacs-template1] **hwtacacs-server traffic-unit mbyte**	（可选）配置 HWTACACS 流量单位。其他说明参见表 17-7 中的第 10 步 不同的 HWTACACS 服务器使用的流量单位可能不同，因此需要在设备上针对每一个 HWTACACS 服务器组设置流量单位，和 HWTACACS 服务器保持一致。**但只有当该 HWTACACS 服务器模板没有用户使用时，才能改变流量单位的配置** 缺省情况下，设备以字节（byte）作为 HWTACACS 流量单位，可用 **undo hwtacacs-server traffic-unit** 命令删除配置的 HWTACACS 流量单位
14	**hwtacacs-server timer response-timeout** *value* 例如：[HUAWEI-hwtacacs-template1] **hwtacacs-server timer response-timeout** 10	（可选）配置 HWTACACS 服务器应答超时时间，取值范围是 1～300 的整数秒。配置超时时间后，设备向 HWTACACS 服务器发出请求报文后，如果在规定的时间内未得到 HWTACACS 服务器发回的应答，需要设备重传请求报文。这样就提高了 HWTACACS 认证、授权、计费过程的可靠性 缺省情况下，HWTACACS 应答超时时间为 5s，可用 **undo hwtacacs-server timer response-timeout** 命令将 HWTACACS 应答超时时间恢复为缺省值

（续表）

步骤	命令	说明
15	**hwtacacs-server timer quiet** *value* 例如：[HUAWEI-hwtacacs-template1] **hwtacacs-server timer quiet** 10	（可选）配置主用服务器恢复激活状态的静默时间，取值范围为 1～255 的整数分钟 【说明】如果主用服务器不可用，设备会自动切换至备用服务器，向备用服务器发送报文。到达主用服务器恢复激活状态的时间后，设备尝试与主用服务器建立连接。 （1）如果主用服务器仍不可用则设备继续向备用服务器发送报文，直到下一次恢复激活状态的时间再次尝试与主用服务器建立连接，如此循环 （2）如果主用服务器可用则设备切换到主用服务器，向主用服务器发送报文 通过设置主用服务器恢复激活状态的静默时间，既保证能够主用服务器尽快恢复激活状态，又减少了服务器切换时的探测次数 缺省情况下，主用服务器恢复激活状态前需要等待 5min，可用 **undo hwtacacs-server timer quiet** 命令恢复主用服务器恢复激活状态的静默时间缺省值
16	**quit** 例如：[HUAWEI-hwtacacs-template1] **quit**	退出 HWTACACS 服务器模板视图，返回系统视图
17	**hwtacacs-server accounting-stop-packet resend** { **disable** | **enable** *number* } 例如：[HUAWEI]**hwtacacs-server accounting-stop-packet resend enable** 50	（可选）配置是否允许重发计费停止报文，以及可重发的计费停止报文个数。命令中的参数和选项说明如下。 （1）**disable**：二选一选项，指定禁止重发计费停止报文，即计费停止报文只发送一次，即使失败了也不会重发 （2）**enable** *number*：二选一参数，指定使能重发计费停止报文，并配置重发的计费停止报文个数，取值范围为 1～300 的整数。如果计费停止报文发送后，收不到回应或者回应失败，会重新发送计费停止报文 缺省情况下，设备启用计费结束报文的重传功能，报文的重传次数为 100，可用 **undo hwtacacs-server accounting-stop-packet resend** 命令恢复计费停止报文重发功能与计费停止报文个数为缺省情况
18	**return** 例如：[HUAWEI] **return**	返回用户视图
19	**hwtacacs-user change-password hwtacacs-server** *template-name* 例如：<HUAWEI> **hwtacacs-user change-password hwtacacs-server** template1	（可选）在设备上修改用户在 HWTACACS 服务器上保存的用户密码。参数 *template-name* 用来指定要修改用户密码的 HWTACACS 模板名称 键入本命令后，系统会给出一个修改用户密码的提示，系统等待超过 30s，用户未输入用户名或新密码、确认密码时，密码修改将中断 【注意】只有在 HWTACACS 服务器上保存的用户名和密码没有过期的情况下，才允许用户主动使用该命令修改密码；对于密码已经过期的用户，登录设备时，HWTACACS 服务器将返回认证不成功，不允许用户主动更改密码。系统允许 HWTACACS 用户修改其他人的密码，当被修改人的权限高于用户本人的权限时，系统允许用户使用 **super** 命令提升用户自身级别后，再使用该命令修改他人密码 用户可以输入 Ctrl+C 取消本次密码修改

【示例 】经过HWTACACS认证的用户主动修改名为huaweiH的WTACACS服务器模板中cj@shy用户的密码。

```
<HUAWEI> hwtacacs-user change-password hwtacacs-server huawei
Info: EXEC is in an interactive process, please wait. .
Username:cj@shy
Old Password:
New Password:
Re-enter New password:
Info: The password has been changed successfully.
```

17.4.3 HWTACACS方式认证、授权和计费配置示例

本示例拓扑结构参见17.3.4节的图17-4。本示例要求采用HWTACACS认证、授权和计费方案，HWTACACS主用服务器为129.7.66.66/24，备用服务器为129.7.66.67/24，服务器的认证、授权和计费端口号均为49。具体用户要求如下。

（1）SwitchB对接入用户先用HWTACACS服务器进行认证，如果认证没有响应，再使用本地认证。

（2）接入的用户进行用户等级提升时，要求先使用HWTACACS对其进行认证，如果HWTACACS认证没有响应，再使用本地认证。

（3）SwitchB对接入用户先用HWTACACS服务器进行授权，如果授权没有响应，再使用本地授权。

（4）SwitchB对接入用户采用HWTACACS计费。

（5）对用户进行实时计费，计费间隔为3min。

1．基本配置思路

根据图17-3（c）所示的配置流程及本示例的具体要求，可得出如下基本配置思路（同样因为示例中没有特定的业务方案要求，所以可不配置业务方案，直接采用缺省配置即可）。

（1）配置AAA方案，包括HWTACACS认证方案、授权方案和计费方案。

（2）配置HWTACACS服务器模板。

（3）在huawei域下应用HWTACACS服务器模板、认证方案、授权方案和计费方案。

以下配置均在SwitchB上进行。

2．具体配置步骤

（1）配置AAA认证方案、授权方案、计费方案。

[HUAWEI] aaa

[HUAWEI-aaa] authentication-scheme l-h　#---配置认证方案l-h

[HUAWEI-aaa-authen-l-h] authentication-mode hwtacacs local　#---配置认证模式为先进行HWTACACS认证，后进行本地认证

[HUAWEI-aaa-authen-l-h] authentication-super hwtacacs super　#---配置用户级别提升认证模式为先进行HWTACACS认证，后进行本地认证

[HUAWEI-aaa-authen-l-h] quit

[HUAWEI-aaa] authorization-scheme hwtacacs　#---配置授权方案hwtacacs

[HUAWEI-aaa-author-hwtacacs] authorization-mode hwtacacs local　#---配置授权模式为先进行HWTACACS授权，后进行本地授权

[HUAWEI-aaa-author-hwtacacs] quit

[HUAWEI-aaa] accounting-scheme hwtacacs　#---配置计费方案hwtacacs

[HUAWEI-aaa-accounting-hwtacacs] accounting-mode hwtacacs　#---配置计费模式为HWTACACS

[HUAWEI-aaa-accounting-hwtacacs] accounting start-fail online　#---配置如果开始计费失败，允许用户上线

[HUAWEI-aaa-accounting-hwtacacs] accounting realtime3　#---配置实时计费间隔为3min

[HUAWEI-aaa-accounting-hwtacacs] quit

（2）配置HWTACACS服务器模板。

[HUAWEI] hwtacacs-server template ht　#---配置HWTACACS服务器模板ht

[HUAWEI-hwtacacs-ht] hwtacacs-server authentication 129.7.66.66 49　#---配置HWTACACS主用认证服务器的IP地址和端口

[HUAWEI-hwtacacs-ht] hwtacacs-server authorization129.7.66.66 49　#---配置HWTACACS主用授权服务器的 IP地址和端口

[HUAWEI-hwtacacs-ht] hwtacacs-server accounting 129.7.66.66 49　#---配置HWTACACS主用计费服务器的IP地址和端口

[HUAWEI-hwtacacs-ht] hwtacacs-server authentication 129.7.66.67 49secondary #---配置HWTACACS备用认证服务器的IP地址和端口

[HUAWEI-hwtacacs-ht] hwtacacs-server authorization 129.7.66.67 49secondary #---配置HWTACACS备用授权服务器的IP地址和端口

[HUAWEI-hwtacacs-ht] hwtacacs-server accounting 129.7.66.67 49secondary #---配置HWTACACS备用计费服务器的IP地址和端口

[HUAWEI-hwtacacs-ht] hwtacacs-server shared-key cipherhello #---配置TACACS服务器共享密钥

[HUAWEI-hwtacacs-ht] quit

（3）配置huawei域，并在huawei域下应用前面配置的HWTACACS认证方案、授权方案、计费方案和HWTACACS服务器模板。

[HUAWEI-aaa] domain huawei

[HUAWEI-aaa-domain-huawei] authentication-scheme l-h

[HUAWEI-aaa-domain-huawei] authorization-scheme hwtacacs

[HUAWEI-aaa-domain-huawei] accounting-scheme hwtacacs

[HUAWEI-aaa-domain-huawei] hwtacacs-serverht

[HUAWEI-aaa-domain-huawei] quit

[HUAWEI-aaa] quit

[HUAWEI] quit

配置好后，在 SwitchB上执行 display hwtacacs-server template命令可以查看到该HWTACACS服务器模板的配置与上面的配置是一致的。具体如下。

<HUAWEI>display hwtacacs-server template ht

```
------------------------------------------------------------------------
HWTACACS-server template name    : ht
Primary-authentication-server    : 129.7.66.66:49:-
Primary-authorization-server     : 129.7.66.66:49:-
Primary-accounting-server        : 129.7.66.66:49:-
Secondary-authentication-server  : 129.7.66.67:49:-
Secondary-authorization-server   : 129.7.66.67:49:-
Secondary-accounting-server      : 129.7.66.67:49:-
Current-authentication-server    : 129.7.66.66:49:-
Current-authorization-server     : 129.7.66.66:49:-
Current-accounting-server        : 129.7.66.66:49:-
Source-IP-address                : 0.0.0.0
Shared-key                       : ****************
Quiet-interval(min)              : 5
Response-timeout-Interval(sec)   : 5
Domain-included                  : Yes
Traffic-unit                     : B
------------------------------------------------------------------------
```

同时可在SwitchB上执行display domain命令查看到该域的配置，具体如下。

```
<HUAWEI> display domain name huawei
Domain-name                    : huawei
Domain-state                   : Active
Authentication-scheme-name     : l-h
Accounting-scheme-name         : hwtacacs
Authorization-scheme-name      : hwtacacs
Service-scheme-name            : -
RADIUS-server-template         : -
HWTACACS-server-template       : ht
User-group                     : -
```

17.5 AAA认证、授权和计费配置管理

全面按照图17-3的流程配置好各种认证、授权和计费方案后，可以使用以下display任意视图命令查看相关配置，使用以下reset用户视图命令清除相关统计信息。

（1）display aaa configuration：查看AAA的摘要配置信息。

（2）display local-user：查看本地用户的摘要信息。

（3）display domain [name domain-name]：查看所有或者指定域的配置信息。

（4）display authentication-scheme [authentication-scheme-name]：查看所有或者指定认证方案的配置信息。

（5）display authorization-scheme [authorization-scheme-name]：查看所有或者指定授权方案的配置信息。

（6）display accounting-scheme [accounting-scheme-name]：查看计费方案的配置信息。

（7）display service-scheme [name name]：查看业务方案的配置信息。

（8）display access-user [domaindomain-name | interface interface-type interface-number [vlanvlan-id [qinqqinq-vlan-id]] |ip-address ip-address [vpn-instancevpn-instance-name] |mac-addressmac-address | slot slot-id | open |user-iduser-number]：查看符合对应条件的所有在线用户的摘要信息。

（9）display radius-server configuration [template template-name]：查看 RADIUS服务器模板的配置信息。

（10）display radius-attribute [template template-name] disable：查看设备所有或者指定模板中禁用的RADIUS属性。

（11）display radius-attribute [template template-name] translate：查看设备所有或者指定模板中RADIUS属性转换的配置信息。

（12）display radius-server accounting-stop-packet { all | ip ip-address }：查看甩有或者指定IP地址的RADIUS服务器的计费停止报文信息。

（13）display hwtacacs-server template [template-name]：查看设备所有或者指定模板中HWTACACS服务器模板的配置信息。

（14）display hwtacacs-server accounting-stop-packet { all | number | ip ip-address }：查看设备所有或者指定IP地址的HWTACACS服务器的计费停止报文信息。

（15）reset aaa {offline-record | abnormal-offline-record | online-fail-record }：清除用户下线、异常下线、

上线失败的记录信息。

（16）reset hwtacacs-server statistics {all |accounting |authentication |authorization }：清除HWTACACS的统计信息。

（17）reset hwtacacs-server accounting-stop-packet { all | ip ip-address }：清除设备所有或者指定IP地址的HWTACACS服务器的计费停止报文统计信息。

（18）reset radius-server accounting-stop-packet { all | ip ip-address }：清除设备所有或者指定IP地址的RADIUS服务器的计费停止报文统计信息。

第18章　NAC配置与管理

在计算机网络安全管理中，用户的网络接入控制（NAC）是必须充分考虑的一件大事，因为现在的网络安全隐患主要不是来自外网，而是内网。而在用户接入控制方面，最直接、最有效的方法就是基于接入设备接口的各种用户认证方法，如本章将要介绍的华为S系列交换机的802.1x认证、MAC认证和Portal认证。它们针对不同的用户需求和实际的网络环境提供的几种基于端口的实用接入控制方案。相对第17章介绍的各种AAA方案来说，此处基于接口的接入控制方法更为直接，直接在接入处进行认证，但同时又不如AAA方案灵活，因为在本章所介绍的这三种认证方式仅是简单的允许或者拒绝接入认证，没有像AAA方案中的为允许接入的用户授予相应的访问权限，更没有为不同用户进行计费的功能。

本章将首先介绍802.1x认证、MAC认证和Portal认证的基本工作原理和基础知识，然后逐个介绍这三种认证方案的各项配置任务的具体配置方法。当然，在其中介绍的这么多功能配置中，绝大多部分是需要根据实际需要和网络环境选择配置的，不必全部同时配置，而且这三种认证中也有许多配置任务是完全或部分相同的。但要注意，在学习本章内容之前，建议先学习一下本书第17章中关于AAA网络访问控制策略方面的基础知识和配置方法，因为这三种认证中所使用的用户账户信息都是在AAA方案下配置的（可以是本地配置的，也可以是在远程的RADIUS服务器上配置的）。

18.1 NAC基础

NAC（Network Admission Control，网络许可控制）是一套从用户终端角度考虑内部网络安全的“端到端”安全解决方案总称，也就是针对用户终端的接入进行严格控制的解决方案。在华为S系列交换机中NAC包括802.1x认证、MAC认证与Portal认证，都是采用图18-1所示的认证模型。它包括用户（User）、网络接入设备（NAD）和接入控制服务器（ACS）三大部分。

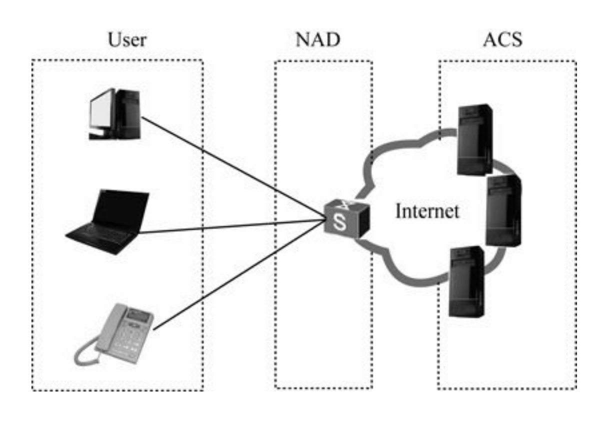


图18-1 NAC认证模型

（1）用户：指接入用户终端，需要对其进行接入认证。如果采用802.1x认证，还需要在用户终端上安装802.1x的客户端软件。

（2）NAD：指网络接入设备（如交换机），对接入用户终端进行认证和授权。一般需要和AAA服务器配合使用（有关AAA的配置与管理参见第17章介绍），防止非法终端接入，防止合法终端越权访问，保护核心资源。

（3）ACS：指接入控制服务器（如 RADIUS 服务器），主要进行终端安全健康性检查与策略管理和用户行为管理与违规审计。本章所介绍的以上三种认证方案中的认证信息（如802.1x认证和Portal认证中的用户名和密码，MAC认证中的用户MAC地址等）都是需要在接入控制服务器中事先配置好的。华为S系列交换机都可以配置RADIUS服务器（具体参见本书第17章），Windows和Linux服务器系统等也可以配置RADIUS服务器。

18.1.1 802.1x认证系统基础

IEEE802.1x是由IEEE制定的关于用户接入网络的认证标准，全称是“基于端口的网络接入控制”。它于2001年正式颁布，最初是为有线网络设计，之后为了配合无线网络的接入进行修订改版，并于2004年完成。

802.1x协议是一种基于端口的网络接入控制协议，所以具体的802.1x认证功能必须在设备端口上进行配置，对端口上接入的用户设备通过认证来控制对网络资源的访问。802.1x认证系统采用网络应用系统典型的Client/Server（C/S）结构，包括3个部分：客户端（Client）、设备端（Device）和认证服务器（Server），如图18-2所示。它与图18-1中的NAC模型结构一一对应。

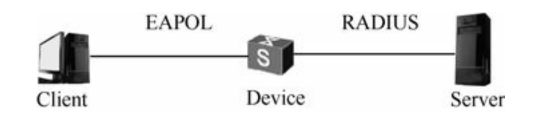


图18-2 802.1x认证系统结构

（1）客户端：局域网用户终端设备，但必须是支持 EAPOL（Extensible Authentication Protocol over LAN，局域网可扩展认证协议）的设备（如 PC 机），可通过启动客户端设备上安装的802.1x客户端软件发起802.1x认证。

（2）设备端：支持 802.1x 协议的网络设备（如交换机），对所连接的客户端进行认证。它为客户端提供接入局域网的端口，可以是物理端口，也可以是逻辑端口（如Eth-Trunk口）。

（3）认证服务器：为设备端802.1x协议提供认证服务的设备，是真正进行认证的设备，实现对用户进行认证、授权和计费，通常为RADIUS服务器。

1．802.1x认证受控/非受控端口

在设备端为客户端提供的接入端口被划分为两个逻辑端口：受控端口和非受控端口。“非受控端口”可看成为 EAP（可扩展认证协议）端口，不进行认证控制，始终处于双向连通状态，主要用来传递在通过认证前必需的 EAPOL 协议帧，保证客户端始终能够发出或接收认证报文。

“受控端口”可以看作为普通业务端口，是需要进行认证控制的。它有“授权”和“非授权”两种状态（相当于在该端口上有一个控制开关）：在授权状态下处于双向连通状态（控制开关闭合），可进行正常的业务报文传递；在非授权状态下处于打开状态（控制开关打开），禁止任何业务报文的传递。设备端利用认证服务器对客户端进行认证的结果（Accept或Reject）来实现对受控端口的授权/非授权状态进行控制。

2．802.1x认证的触发方式

在华为S系列交换机中，802.1x的认证过程可以由客户端主动发起，也可以由设备端主动发起。在“客户端主动触发方式”中，由客户端主动向设备端发送 EAPOL-Start （EAPOL开始）报文来触发认证；而“设备

端主动触发方式”中用于支持不能主动发送EAPOL-Start报文的客户端，例如Windows XP自带的 802.1x客户端。

在“设备端主动触发方式”中又有两种以下具体的触发方式。

（1）DHCP报文触发：设备在收到用户的DHCP请求报文后主动触发对用户的802.1x认证，仅适用于客户端采用**DHCP**方式自动分配**IP**地址的情形。因为**DHCP**请求报文是以广播方式发送的，所以在同一网段中的设备都可以收到，故设备端不一定就是担当**DHCP**服务器的设备 。

（2）源 MAC 地址未知报文触发：当设备收到源 MAC 地址未知的报文时主动触发对用户的802.1x认证。若设备端在设置的时长内没有收到客户端的响应，则重发该报文。

3．802.1x的认证方式

无论是哪种触发方式，802.1x认证系统都是使用EAP协议来实现客户端、设备端和认证服务器之间认证信息的交换。在客户端与设备端之间使用的是基于以太局域网的EAPOL格式封装EAP报文，然后承载于以太网数据帧中进行交互；而设备端与RADIUS服务器之间的EAP报文可以使用以下两种方式进行交互。

（1）EAP中继：来自客户端的EAP报文到达设备端后，直接使用EAPOR（EAP over RADIUS）格式封装在RADIUS报文中，再发送给RADIUS服务器，则RADIUS服务器 从封装的EAP报文中获取客户端认证信息，然后对客户端进行认证。

这种认证方式的优点是设备端的工作很简单，不需要对来自客户端的EAP报文进行任何处理，只需要用EAPOR对EAP报文进行封装即可，根本不管客户端的认证信息。同时在这种认证方式中，设备端与RADIUS 服务器之间可支持多种 EAP 认证方法，例如MD5-Challenge、EAP-TLS、PEAP等，但要求服务器端也支持相应的认证方法。

（2）EAP终结：来自客户端的EAP报文在设备端进行终结，然后由设备端将从EAP报文中提取的客户端认证信息封装在标准的RADIUS报文（不再是EAPOR格式）中，与RADIUS服务器之间采用PAP（Password Authentication Protocol，密码验证协议）或CHAP（Challenge Handshake Authentication Protocal，质询握手验证协议）方式对客户端进行认证（当然在RAIUDS服务器端必须配置合法用户的用户名和密码信息）。

这种认证方式的优点是现有的RADIUS服务器基本均可支持PAP和CHAP认证，无需升级服务器，但设备端的工作比较繁重，因为在这种认证方式中，设备端不仅要从来自客户端的 EAP 报文中提取客户端认证信息，还要通过标准的 RAIUDS 协议对这些信息进行封装，且不能支持除MD5-Challenge之外的其他EAP认证方法。

4．802.1x认证支持的Guest VLAN、Restrict VLAN与Critical VLAN

为了使那些不支持 802.1x 的客户端通过安装或者升级 802.1x 客户端软件来支持802.1x认证，也为那些在认证过程中认证失败，或认证服务器无响应时提供一些基本的访问资源，华为S系列交换机的802.1x认证功能提供了以下三种不同的特殊VLAN功能（当然，这些都是可选的配置 ）。

（1）Guest VLAN。Guest VLAN（来宾VLAN）功能开启后，当用户不响应 **802.1x**认证请求 时（如未安装客户端软件），设备端会将用户所在端口加入到Guest VLAN中，以便这些用户可以访问 Guest VLAN，获取客户端软件，升级客户端或执行其他一些用户升级程序等操作。

（2）Restrict VLAN。Restrict VLAN（限制VLAN）功能开启后，当用户认证失败 时（如输入了错误的用户名和密码），设备端会将用户所在端口加入到Restrict VLAN中。Restrict VLAN和Guest VLAN的功能相似，都是满足用户在通过认证前可以访问有限的网络资源。但通常在Restrict VLAN中部署的网络资源比Guest VLAN中更少，从而更严格地限制未通过认证的用户对网络资源的访问。

（3）Critical VLAN。Critical VLAN（严格VLAN）功能开启后，当认证服务器无响应 时（如设备端与

认证服务器之间的网络断开或者认证服务器出现故障），设备端会将用户所在端口加入Critical VLAN中进而能够访问Critical VLAN中的资源。

5．802.1x快速部署

在实际的应用中，如果网络规模比较大，且有大量不能很好支持802.1x功能的客户端，这时客户端的部署工作量可能很大。为此华为S系列交换机提供了802.1x认证快速部署功能。它可引导这些用户自助下载安装客户端，实现客户端的快速部署。

802.1x认证快速部署功能通过以下两个功能实现。

（1）用户受限访问。802.1x 认证成功之前，通过 ACL 限制终端用户只能访问一个特定的IP地址段（免认证IP网段），或是特定服务器，在特定服务器上提供客户端的下载升级或者动态地址分配等服务。

（2）用户HTTP访问URL重定向功能。用户在未进行802.1x认证前或者认证失败后，使用浏览器访问网络时，设备会将用户访问的URL重定向到配置好的URL，供这些用户从这个网页中下载客户端软件。

6．用户组授权功能

设备支持根据用户组对用户进行授权控制（就相当于 Windows 系统中的用户组一样）。在用户认证成功后，认证服务器为该用户下发到对应的用户组，使该用户具备该用户组中的访问权限。每个用户组可以关联不同的ACL规则，通过用户组和ACL规则的关联，实现对每类用户进行 ACL 授权信息控制，即同类用户采用同样的授权信息。

18.1.2 802.1x认证原理

前面说了，在802.1x认证过程中，设备端与RADIUS服务器之间支持EAP中继和EAP终结两种认证方式。下面均以客户端主动发起认证 为例介绍这两种认证方式的工作原理。

1．EAP中继认证原理

在EAP中继认证的过程中，设备端起一个中继代理的角色，用于通过EAPOR封装和解封装的过程转发客户端和认证服务器之间的交互报文。整个认证过程是先进行用户名认证，再进行对应的密码认证，具体如图18-3所示（对应图18-3中的序号）。

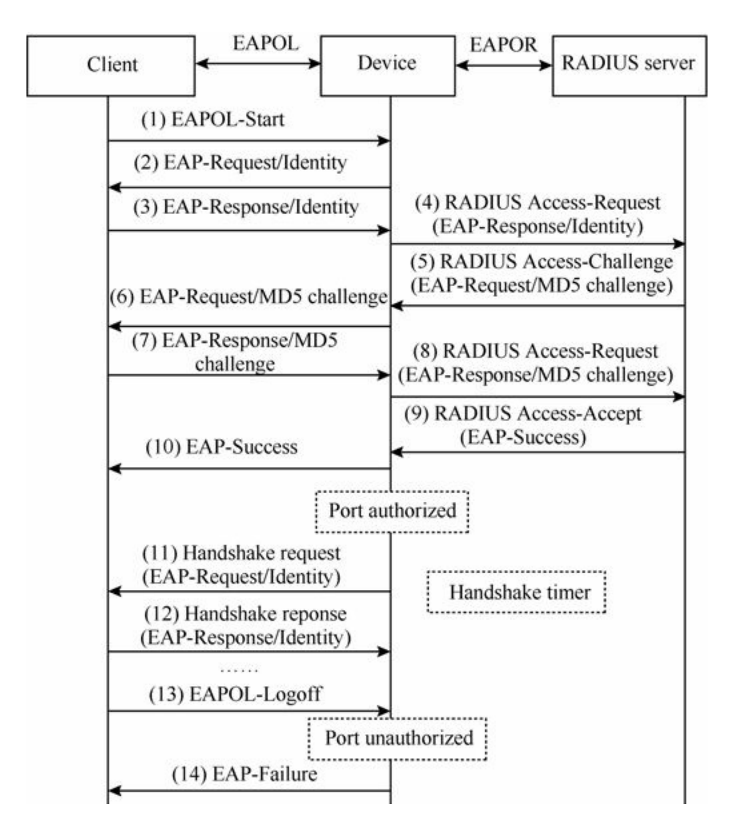


图18-3 EAP中继认证流程

（1）当用户访问网络时自动打开802.1x客户端程序，根据提示输入已经在RADIUS服务器中创建的用户名和密码，发起连接请求。因为端口最初的状态是未授权状态，所 以此时端口除了 IEEE 802.1x协议包外不能接收和发送任何包。此时，客户端程序将向设备端发出认证请求帧（EAPOL-Start），启动认证过程。

（2）设备端在收到客户端的认证请求帧后，将发出一个 Identity（标识）类型的EAP 请求帧（EAP-Request/Identity），要求用户的客户端程序发送上一步用户所输入的用户名。

（3）客户端程序在收到设备端的Identity请求帧后，将用户名信息通过Identity类型的EAP响应帧（EAP-Response/Identity）发送给设备端，响应设备端发出的请求。

（4）设备端将客户端发送的Identity响应帧中的EAP报文原封不动地使用EAPOR格式封装在RADIUS报文（ RADIUS Access-Request）中，发送给认证服务器进行处理。

（5）RADIUS服务器收到设备端发来的RADIUS报文后从中提取用户名信息后，将该信息与数据库中的用户名列表中对比，找到该用户名对应的密码信息，并用随机生成的一个MD5 Challenge消息对密码进行加密处理，然后将此MD5 Challenge消息同样通过EAPOR格式封装以RADIUS Access-Challenge报文发送给设备端。

（6）设备端在收到来自RADIUS服务器的EAPOR格式的Access-Challenge报文后，通过解封装，将其中的MD5 Challenge消息转发给客户端。

（7）客户端在收到由设备端传来的MD5 Challenge消息后，用该Challenge消息对密码部分进行加密处理，然后生成 EAP-Response/MD5 Challenge报文，并发送给设备端。

（8）设备端又将此EAP-Response/MD5 Challenge报文以EAPOR格式封装在RADIUS报文（RADIUS Access-Request）中发送给RADIUS服务器。

（9）RADIUS服务器收到的已加密的密码信息后，与第（5）步在本地经过加密运算后的密码信息进行

对比，如果相同则认为是合法用户，并向设备端发送认证通过报文（RADIUS Access-Accept）。

（10）设备收到RADIUS Access-Accept报文后，经过EAPOR解封装再以EAP-Success报文向客户端发送，并将端口改为授权状态，允许用户通过端口访问网络。

（11）用户在线期间设备端会通过向客户端定期发送握手报文，对用户的在线情况进行监测。

（12）客户端收到握手报文后向设备发送应答报文，表示用户仍然在线。缺省情况下，若设备端发送的两次握手请求报文都未得到客户端应答，设备端就会让用户下线，防止用户因为异常原因下线而设备无法感知。

（13）客户端可以发送EAPOL-Logoff帧给设备端，主动要求下线。

（14）在设备端收到客户端发来的EAPOL-Logoff帧后，把端口状态从授权状态改变成未授权状态，并向客户端发送EAP-Failure报文，确认对应客户端下线。

2．EAP终结认证原理

EAP终结方式与EAP中继方式的认证流程相比，主要不同在于步骤（4）中用来对用户密码信息进行加密处理的MD5 challenge是由设备端生成（而不是由RADIUS服务器生成），之后设备端会把用户名、MD5 challenge和客户端加密后的密码信息一起送给RADIUS服务器，进行相关的认证处理。具体流程如图18-4所示。

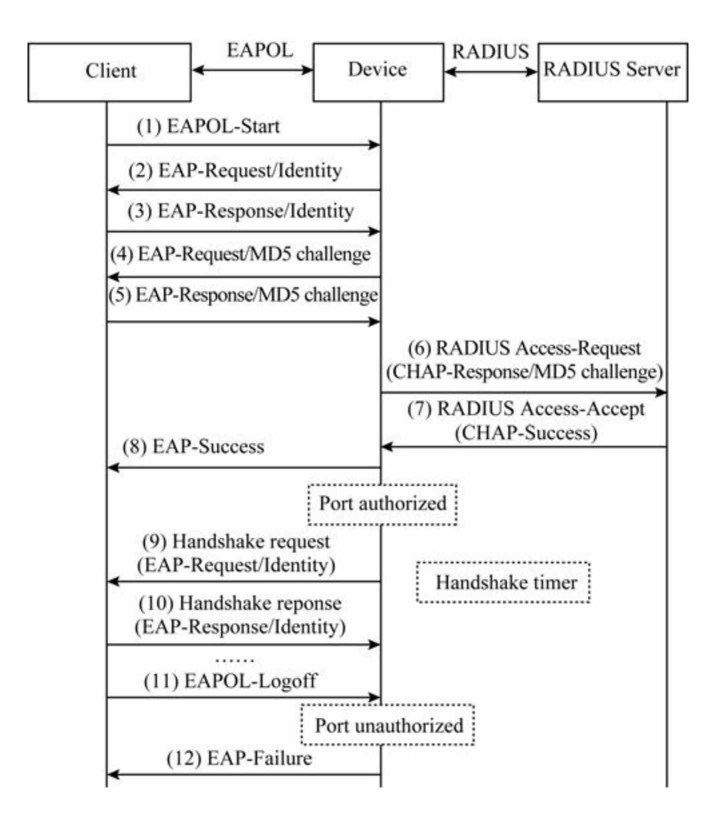


图18-4 EAP终结认证流程

3．MAC旁路认证

在 802.1x 认证过程中，设备端会首先触发用户采用 802.1x 认证方式，但若用户长时间内没有进行802.1x认证（如图18-5所示），则以用户的MAC地址作为用户名和密码上送认证服务器进行认证。MAC旁路认证可使802.1x认证系统中无法安装和使用802.1x客户端软件的终端，例如打印机等，以自身MAC地址作为用户名和密码进行认证。

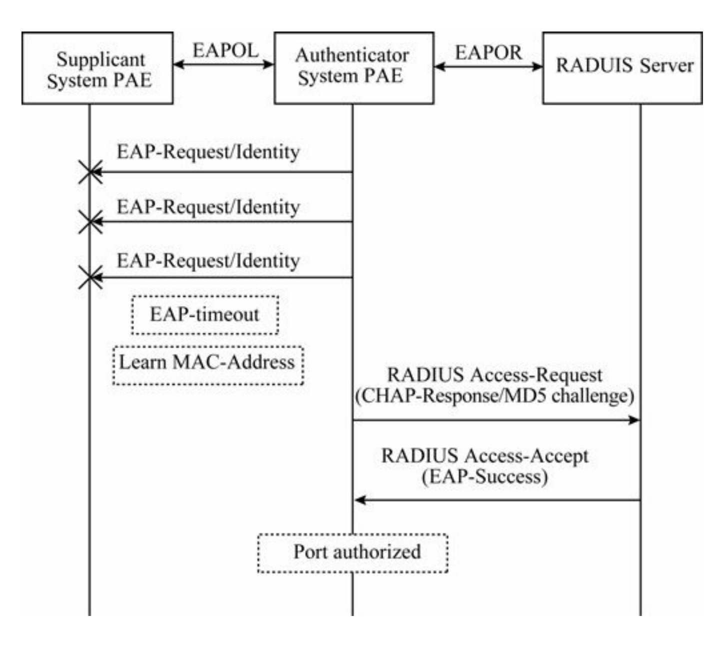


图18-5 MAC旁路认证流程

18.1.3 MAC认证

MAC 认证是一种基于端口和 MAC 地址对用户的网络访问权限进行控制的认证方法。它与前面介绍的802.1x认证相比，最大的特点就是不需要用户安装任何客户端软件，也不需要用户手动输入用户名或者密码，设备在启动了 MAC 认证的端口上首次检测到用户的 MAC 地址以后即启动对该用户的认证操作。很显然这不是一种十分安全的认证方式，因为 MAC 地址完全可以仿冒。但 MAC 认证确实是一种比较有用的认证方式，因为它的配置和认证过程都很简单。

根据设备端最终用于验证用户身份的用户名格式和内容的不同，可以将 MAC 认证使用的用户名格式分为两种类型。

（1）MAC地址用户名格式：使用用户的MAC地址作为认证时的用户名和密码，适用于一个端口连接一个客户端的情形。

（2）固定用户名形式：不论用户的 MAC 地址为何值，所有用户均使用设备上指定的一个固定用户名和密码替代用户的 MAC 地址作为身份信息进行认证。适用于同一个端口下连接多个用户的情形，但要求接入的客户端比较可信。

MAC认证与前面介绍的 802.1x认证一样，也支持“Guest VLAN功能”和“用户组授权功能”。在Guest VLAN功能开启后，当用户不响应MAC认证请求时，设备会将用户所在端口加入Guest VLAN中，这样用户就可以访问Guest VLAN中的资源，从而满足了用户不进行认证也能够访问某些基本的网络资源的需求。

在用户组授权功能开启后，设备支持根据用户组对用户进行授权控制，即用户认证成功后，认证服务器下发用户组，将用户进行分类。每个用户组可以关联不同的 ACL规则，通过用户组和ACL规则的关联，实现对每类用户进行ACL授权信息控制，即同类用户采用同样的授权信息。

18.1.4 Portal认证

Portal认证通常也称为Web认证，是通过Web页面进行认证的，一般将用于Portal认证的网站称为门户网站。未认证用户上网时，设备强制用户登录到特定站点，用户可以免费访问其中的服务。当用户需要使用

互联网中的其他信息时，必须在门户网站进行认证，只有认证通过后才可以使用互联网资源。

用户可以主动访问已知的 Portal 认证网站，输入用户名和密码进行认证，这种开始Portal认证的方式称作主动认证。反之，如果用户试图通过HTTP访问其他外网，将被强制访问 Portal认证网站，从而开始 Portal认证过程，这种方式称作强制认证。

1．Portal认证系统结构

Portal服务器可以是接入设备之外的独立设备担当，称之为“外置Portal服务器”，也可以由接入设备自己担当，称之为“内置 Portal 服务器”。使用外置 Portal 服务器的Portal 认证系统由 4 个基本要素组成：认证客户端、接入设备、Portal 服务器与认证/计费服务器，如图18-6所示。

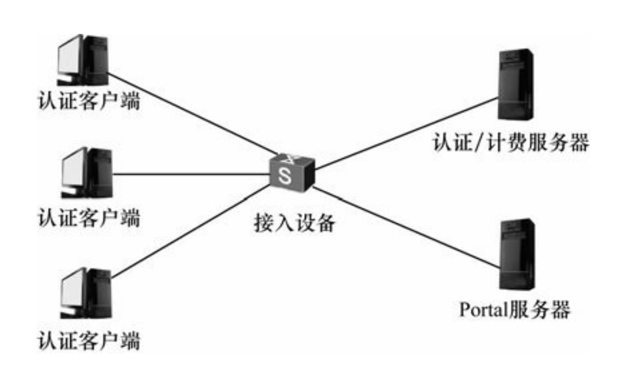


图18-6 使用外置Portal服务器的Portal认证系统组成

（1）认证客户端：运行HTTP/HTTPS协议的浏览器或运行Portal客户端软件的主机。

（2）接入设备：与认证客户端连接的设备（如交换机、路由器等），提供以下三方面作用。

① 在认证之前，将认证网段内用户的所有HTTP请求都重定向到Portal服务器。

② 在认证过程中，与Portal服务器、认证/计费服务器交互，完成身份认证/计费的功能。

③ 在认证通过后，允许用户访问被管理员授权的互联网资源。

（3）Portal 服务器：接收 Portal 客户端认证请求的服务器端系统，提供免费门户服务和基于Web认证的界面，与接入设备交互认证客户端的认证信息。

（4）认证/计费服务器：与接入设备进行交互，完成对用户的认证和计费。

虽然看起来有4个部分，其实仍与前面图18-1所说到的NAC模型的3个部分是一样的，因为Portal服务器和认证/计费服务器都属于ACS。

使用内置Portal服务器的Portal认证系统由3个基本要素组成：认证客户端、接入设备和认证/计费服务器，如图18-7所示。这时，不同的只是接入设备与Portal服务器是由一个设备担当。



图18-7 使用内置Portal服务器的Portal认证系统组

通过内置Portal服务器进行Portal认证，由于不需要部署额外的Portal服务器，故增强了Portal认证的通用性。但通常，同时担当内置Portal服务器的接入设备提供了比较简单的Portal服务器功能，仅能给用户提供通过Web方式上线、下线的基本功能，并不能完全替代独立的Portal服务器，也不支持外置独立服务器的任何扩展功能，例如二次地址分配等。

2．Portal认证方式

不同的组网方式下，可采用的Portal认证方式不同。按照网络中实施Portal认证的网络层次来分，Portal的认证方式分为两种：二层认证方式和三层认证方式。

（1）二层认证方式

当认证客户端与接入设备直连（或之间只有二层设备存在）时，设备能够学习到用 户的MAC地址，则设备可以利用IP和MAC地址来识别用户，此时可配置Portal认证为二层认证方式。

二层认证流程简单，但由于限制了用户只能与接入设备处于同一网段，降低了组网的灵活性。二层认证方式的报文交互过程如图18-8所示，具体描述如下（流程号对应图中的序号）。

① 认证客户端进行HTTP访问时，HTTP报文进入接入设备时，接入设备要进行分辨：对于访问Portal服务器或设定的免认证网络资源的HTTP报文，允许其通过；对于访问其他地址的HTTP报文，将HTTP访问重定向到Portal服务器。Portal服务器提供Web页面供用户输入用户名和密码来进行认证。

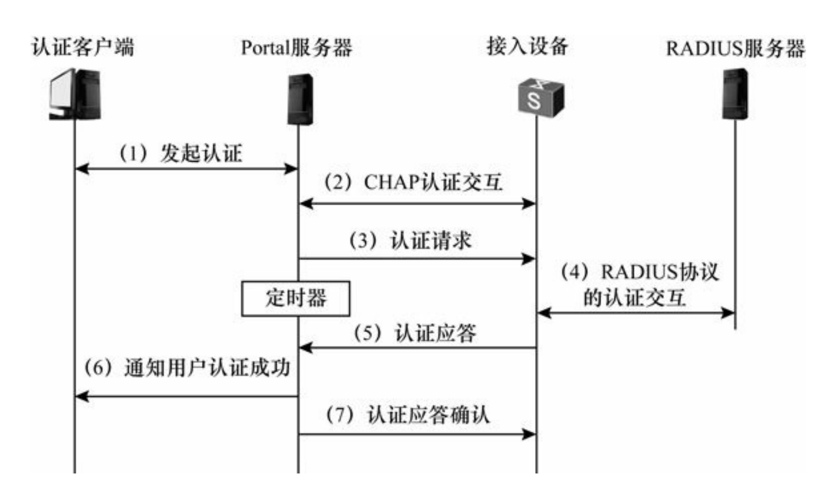


图18-8 Portal二层认证流程

② Portal服务器与接入设备之间进行CHAP认证交互。如果采用PAP认证，则Portal服务器无需与接入设备进行PAP认证交互，不进行本步流程，直接进行下面的第③步。

③ Portal服务器将用户输入的用户名和密码组装成认证请求报文发往接入设备，同时开启定时器等待认证应答报文。

④ 接入设备与RADIUS服务器之间进行RADIUS协议报文的交互。

⑤ 接入设备向Portal服务器发送认证应答报文。

⑥ Portal服务器向客户端发送认证通过报文，通知客户端认证成功。

⑦ Portal服务器向接入设备发送认证应答确认。

（2）三层认证方式

当设备部署在汇聚层或核心层时，在认证客户端和设备之间存在三层转发设备，此时设备不一定能获取到认证客户端的MAC地址，所以将以IP地址唯一标识用户，此时需要将Portal认证配置为三层认证方式。

三层认证的报文处理流程跟二层认证完全一致。三层认证组网灵活，容易实现远程控制，但由于只有IP可以用来标识一个用户，所以安全性不高。

3．Portal认证探测与逃生功能

在Portal认证实际组网应用中，如果接入设备与Portal服务器之间出现网络故障，导致通信中断或者Portal服务器本身出现故障，就会造成新的Portal认证用户无法上线，已经在线的Portal用户也无法正常下

线。这时就要用到Portal认证的探测和逃生功能， 它可使在网络故障或Portal服务器无法正常工作的情况下，接入设备让用户仍然能够正常使用网络，并具有一定的网络访问权限，并通过日志和Trap的方式报告故障。同时在故障修复后，接入设备通过用户信息同步机制可保证Portal服务器与设备上用户信息的一致性，以避免可能出现的计费不准确问题。

4．用户组授权功能

与801.1x认证和MAC认证一样，Portal认证也支持根据用户组对用户进行授权控制。即用户认证成功后，认证服务器下发用户组，将用户进行分类。每个用户组可以关联不同的ACL规则，通过用户组和ACL规则的关联，实现对每类用户进行ACL授权信息控制，即同类用户采用同样的授权信息。

18.1.5 NAC特性的产品支持

华为S系列交换机同时支持802.1x认证、MAC认证、Portal认证以及混合认证。但各种不同系列所支持的这些认证特性不完全相同（除非特别说明，所有S系列交换机均支持所列的各种认证特性），下面分别予以介绍。

1．802.1x认证特性及产品支持

除了基本的802.1x认证功能外，为了使管理员能够更加合理有序地控制、管理802.1x用户，华为S系列交换机还支持以下802.1x认证特性。

（1）支持MAC旁路认证功能。

（2）支持对多个定时器的值进行设置。

（3）支持Guest VLAN、Restrict VLAN与Critical VLAN功能。但**S2700/3700**系列不支持**Restrict VLAN**与**Critical VLAN**功能 。

（4）支持通过DHCP报文或源MAC未知报文触发802.1x认证。但**S2700/3700**系列不支持源**MAC**未知报文触发**802.1x**认证 。

（5）支持802.1x认证快速部署功能，但**S2700/3700**系列不支持 。

2．MAC认证特性及产品支持

除了基本的MAC认证功能外，为了使管理员能够更加合理有序地控制、管理MAC认证用户，华为S系列交换机还支持以下MAC认证特性。

（1）支持用户名形式以及用户认证域的设置。

（2）支持配置接口允许接入的最大MAC认证用户数。

（3）支持对多个定时器的值进行设置。

（4）支持对MAC认证用户进行重认证功能。

（5）支持Guest VLAN功能。

3．Portal认证及产品支持

除了基本的Portal认证功能外（**S5700LI**和**S5700S-LI**子系列不支持**Portal**认证 ），为了使管理员能够更加合理有序地控制、管理设备与Portal服务器的信息交互以及Portal认证用户，华为S系列交换机还支持以下Portal认证特性。

（1）支持采用外置Portal服务器与内置Portal服务器的Portal认证。但**S2700/3700**系列，以及**S7700/9300/9300E/9700**系列均不支持内置**Portal**服务器的**Portal**认证 。

（2）支持配置设备与Portal服务器信息交互参数。

（3）支持配置Portal认证用户接入控制参数。

（4）支持配置Portal认证用户下线探测周期。但**S2700/3700**系列不支持 。

（5）支持Portal认证的探测与逃生功能。但**S2700/3700**系列不支持 。

4．混合认证

为了灵活地适应网络环境中的多种认证需求，一些华为S系列交换机支持在接入用户的端口上同时配置802.1x认证、MAC认证、Portal认证，使得用户可以选择任何一种适合的认证机制来进行认证，且只需要成功通过一种方式的认证即可实现接入。

但**S2700/3700**系列，以及**S7700/9300/9300E/9700**系列不支持混合认证配置，在支持的混合认证的**S5700/6700**系列中，**Portal**认证也仅支持使用内置**Portal**服务器认证场景，且仅**S5700EI**、**S5700HI**、**S5710EI**子系列和**S6700**系列支持混合认证配置 。

5．三种认证方式比较

以上介绍的802.1x认证、Portal认证和MAC认证的比较如表18-1所示。

表18-1 三种NAC认证方式比较

对比项	802.1x 认证	Portal 认证	MAC 认证
客户端程序需求	需要	Portal 需要，Web 强推不需要	不需要
优点	部署在接入层时，直接控制网络接入口的通断，安全性高	部署灵活	无需安装客户端
缺点	部署不灵活	安全性不高	管理复杂，需登记 MAC 地址
适合场景	新建网络，用户集中，且信息安全要求严格的场景	认证方式灵活，适用于用户分散场景	适用于打印机，传真机等哑终端接入认证的场景

18.1.6 各种NAC认证方式的缺省配置

华为S系列交换机的NAC中的802.1x认证、MAC认证和Portal认证都有缺省配置，分别如表18-2、表18-3和表18-4所示。实际应用的配置可以基于缺省配置进行修改。

表18-2 802.1x认证的缺省配置

参数	缺省值
802.1 认证	未使能
接口授权状态	自动识别模式（auto）
接口接入控制方式	基于 MAC 方式
用户认证方式	CHAP 认证
握手定时器（handshake-period）	S2700/3700/5700/6700 系列为 15s，S700/9300/9300E/9700 系列为 60s
静默定时器（quiet-period）	60s
周期性重认证定时器（reauthenticate-period）	3600s
认证服务器超时定时器（server-timeout）	30s
客户端认证超时定时器（client-timeout）	30s
用户名请求超时定时器（tx-period）	30s
向用户发送认证请求报文的最大次数	2 次

表18-3 MAC认证的缺省配置

参数	缺省值
MAC 认证	未使能
用户名形式	MAC 认证的用户名和密码均为**不带分隔符“-”**的 MAC 地址
用户认证域	default
Guest-Vlan 用户重认证定时器（guest-vlan reauthenticate-period）	60s
用于下线探测定时器（offline-detect）	300s
静默定时器（quiet-period）	60s
周期性重认证定时器（reauthenticate-period）	1 800s
认证服务器超时定时器（server-timeout）	30s

表18-4 Portal认证的缺省配置

参数	缺省值
Portal 认证	未使能
设备支持的 Portal 协议版本	v2、v1
设备向 Portal 服务器发送报文时使用的目的端口号	50100
设备侦听 Portal 协议报文的端口号	2000
Portal 认证的源认证网段	0.0.0.0/0
内置 Portal 服务器对 Portal 认证用户的认证方式	CHAP 方式
下线探测周期	300s

18.2 802.1x认证配置与管理

通过配置802.1x认证可以实现基于接口的用户接入控制，但802.1x认证只提供了一个用户身份认证的实现方案，为了完成用户的身份认证还需要选择使用RADIUS或本地认证方法。因此，需要首先完成以下配置任务（具体请参见本书第17章）。

（1）配置用户所属的ISP认证域及其使用的AAA方案，即本地认证方案或RADIUS方案。

（2）如果需要通过RADIUS服务器进行认证，则应该在RADIUS服务器上配置相应的用户名和密码；如果需要本地认证，则应该在网络接入设备上手动添加认证的用户名和密码。

802.1x认证可配置的任务如下（仅第一项为必选的，其余均为可选配置任务 ）。

（1）使能802.1x认证功能。

（2）（可选）配置接口授权状态。

（3）（可选）配置接口接入控制方式。

（4）（可选）配置用户认证方式。

（5）（可选）使能MAC旁路认证功能。

（6）（可选）配置接口允许接入的最大802.1x认证用户数。

（7）（可选）配置802.1x认证的定时器。

（8）（可选）配置802.1x认证的静默功能。

（9）（可选）配置对802.1x认证用户进行重认证。

（10）（可选）配置802.1x在线用户握手功能。

（11）（可选）配置Guest VLAN功能。

（12）（可选）配置Restrict VLAN功能。

（13）（可选）配置Critical VLAN功能。

（14）（可选）配置802.1x认证的接口Open功能。

（15）（可选）配置允许DHCP报文触发802.1x认证。

（16）（可选）配置单播报文触发802.1x认证。

（17）（可选）配置802.1x快速部署功能。

（18）（可选）配置用户组功能。

下面各小节分别介绍以上各项配置任务。

18.2.1 使能802.1x认证功能

只有同时使能全局和接口上的802.1x认证功能，802.1x的其他配置才能在接口上生效，具体配置步骤如表 18-5 所示。但如果在接口下有 **802.1x** 在线用户，则不能去使能**802.1x**认证功能 。

表18-5 使能802.1x认证功能的配置步骤

步骤	命令	说明		
1	**system-view** 例如：<HUAWEI> **system-view**	进入系统视图		
2	**dot1x enable** 例如：[HUAWEI] **dot1x enable**	使能全局 802.1x 认证功能。缺省情况下，未使能全局802.1x 认证功能，可用 **undo dot1x enable** 命令去使能设备的全局 802.1x 认证功能		
3	**dot1x enable interface** { *interface-type interface-number1* [**to** *interface-number2*] } &<1-10> 例如：[HUAWEI]**dot1x enable interface** gigabitethernet 0/0/1 **to** 0/0/4	在系统视图下批量为多个接口使能802.1x 认证功能。命令中的参数说明如下。 （1）{ *interface-type interface-number1* [**to** *interface-number2*] }：指定要使能802.1x 认证功能的一个接口或者一个连续范围的多个接口 （2）&<1-10>：表示前面的{ *interface-type interface-number1* [**to** *interface-number2*] }参数最多可以有 10 个 缺省情况下，所有接口都没有使能802.1x认证功能，可用**undo dot1x enable** [**interface** { *interface-type interface-number1* [**to** *interface-number2*] } &<1-10>]命令去使能设备指定接口上的 802.1x 认证功能	在系统视图下批量使能多个接口的802.1x 认证功能	（二选一）

（续表）

步骤	命令	说明		
4	**interface** *interface-type interface-number* 例如：[HUAWEI]**interface** gigabitethernet 0/0/1	键入要使能 802.1x 认证功能的接口，进入接口视图	在具体接口视图下使能802.1x 认证功能	（二选一）
5	**dot1x enable** 例如：[HUAWEI-GigabitEthernet0/0/1]**dot1x enable**	使能以上接口的 802.1x 认证功能 缺省情况下，未使能接口的 802.1x 认证功能，可用 **undo dot1x enable** 命令去使能以上接口的 802.1x 认证功能		

18.2.2 （可选）配置接口授权状态

通过配置接口的授权状态，可以控制接入用户是否需要经过认证来访问网络资源。华为S系列交换机接口支持3种802.1x认证授权状态。

（1）自动识别模式（auto）：接口初始状态为非授权状态，仅允许收发EAPOL报文，不允许用户访问网络资源；认证通过后接口切换到授权状态，允许用户访问网络资源。

（2）强制授权模式（authorized-force）：接口始终处于授权状态，允许用户不经认证授权即可访问网络资源。仅适用于信任客户端所连接的接口 。

（3）强制非授权模式（unauthorized-force）：接口始终处于非授权状态，不允许用户访问网络资源。仅适用于非信任客户端所连接的接口 。

接口授权状态可在系统视图下为多个接口进行批量配置，或在接口视图下为单个接口配置，但都是针对接口进行配置的。具体步骤如表18-6所示。

表18-6 接口授权状态的配置步骤

步骤	命令	说明	
1	**system-view** 例如：<HUAWEI> **system-view**	进入系统视图	
2	**dot1x port-control** { **auto** | **authorized-force** | **unauthorized-force** } **interface** { *interface-type interface-number1*[**to** *interface-number2*] } &<1-10> 例如：[HUAWEI] **dot1x port-control authorized-force interface** gigabitethernet 0/0/1 **to** 0/0/4	在系统视图下批量为多个接口配置授权状态。命令中的参数和选项说明如下。 （1）**auto**：多选一选项，指定接口的授权状态为自动识别。此时，接口初始状态为非授权状态，仅允许收发 EAPOL 报文，不允许用户访问网络资源；认证通过后接口切换到授权状态，允许用户访问网络资源 （2）**authorized-force**：多选一选项，指定接口的授权状态为强制授权模式。此时，接口始终处于授权状态，允许用户不经认证授权即可访问网络资源 （3）**unauthorized-force**：多选一选项，指定接口的授权状态为强制非授权模式。此时，接口始终处于非授权状态，不允许用户访问网络资源 （4）{ *interface-type interface-number1* [**to** *interface- number2*] }：指定要配置授权状态的一个接口或者一个连续范围的多个接口	在系统视图下为一个或多个接口批量配置接口的授权状态
2	**dot1x port-control** { **auto** | **authorized-force** | **unauthorized-force** } **interface** { *interface-type interface-number1*[**to** *interface-number2*] } &<1-10> 例如：[HUAWEI] **dot1x port-control authorized-force interface** gigabitethernet 0/0/1 **to** 0/0/4	（5）&<1-10>：表示前面的 { *interface-type interface-number1* [**to** *interface-number2*] } 参数最多可以有 10 个 缺省情况下，接口的授权状态为 **auto**，可用 **undo dot1x port-control interface** { *interface-type interface- number1* [**to** *interface-number2*] } &<1-10>命令恢复指定接口的授权状态为缺省状态	在系统视图下为一个或多个接口批量配置接口的授权状态
3	**interface** *interface-type interface-number* 例如：[HUAWEI]**interface** gigabitethernet 0/0/1	键入要配置接口授权状态的接口，进入接口视图	在具体接口视图下为单个接口配置授权状态
4	**dot1x port-control** { **auto** | **authorized-force** | **unauthorized-force** } 例如：[HUAWEI-GigabitEthernet0/0/1] **dot1x port-control authorized-force**	配置以上接口的授权状态。命令中的选项参见前面第 2 步说明。缺省情况下，接口的授权状态为 **auto**，可用 **undo dot1x port-control** 命令恢复接口的授权状态为缺省的状态	

18.2.3 （可选）配置接口接入控制方式

在使能 802.1x 认证功能后，设备支持两种接口接入控制方式（当有 **802.1x** 用户在线时，不允许更改对应的接口接入控制方式 ）。

（1）基于接口方式：在这种接入控制方式下，该接口下的第一个用户认证成功后，其他接入用户无需认证即可直接接入网络。但同时，当该认证成功的用户下线后，其他用户也将被拒绝接入网络。此接入控制方式适合于均需要通过认证的集团用户。

（2）基于 MAC 地址方式：接口下的所有接入用户均需单独认证，当某个用户下线时，不影响其他用

户接入网络。此接入控制方式比较适合零散用户。

接口接入控制方式可在系统视图下为多个接口进行批量配置，或在具体的接口视图下为单个接口配置，具体配置步骤如表18-7所示。

表18-7 接口接入控制方式的配置步骤

步骤	命令	说明	
1	**system-view** 例如：<HUAWEI> **system-view**	进入系统视图	
2	**dot1x port-method** { **mac** | **port** } **interface** { *interface-type interface-number1* [**to** *interface-number2*] } &<1-10> 例如：[HUAWEI]**dot1x port-method port interface** gigabitethernet 0/0/1 **to** 0/0/4	在系统视图下批量为多个接口配置接入控制方式。命令中的参数和选项说明如下。 （1）**mac**：二选一选项，指定基于 MAC 地址的接入控制方式，即采用用户的 MAC 地址对用户进行认证 （2）**port**：二选一选项，指定基于接口的接入控制方式，即采用用户所连接的交换机接口对用户进行认证 （3）{ *interface-type interface-number1* [**to** *interface-number2*] }：指定要配置接入控制方式的一个接口或者一个连续范围的多个接口 （4）&<1-10>：表示前面的{ *interface-type interface-number1* [**to** *interface-number2*] } 参数最多可以有 10 个 缺省情况下，802.1x 在指定接口上进行接入控制的方式为基于 MAC 地址方式，可用 **undo dot1x port-method interface** { *interface-type interface-number1* [**to** *interface-number2*] } &<1-10>命令恢复 802.1x 在指定接口上的接入控制方式为缺省方式	在系统视图下为一个或多个接口批量配置接入控制方式
3	**interface** *interface-type interface-number* 例如：[HUAWEI]**interface** gigabitethernet 0/0/1	键入要配置接入控制方式的接口，进入接口视图	在具体接口视图下为单个接口配置接入控制方式
4	**dot1x port-method** { **mac** | **port** } 例如：[HUAWEI-GigabitEthernet0/0/1]**dot1x port-method port**	配置以上接口的接入控制方式。命令中的选项参见前面第 2 步说明 缺省情况下，802.1x 在指定接口上进行接入控制的方式为基于 MAC 地址方式，可用 **undo dot1x port-method interface** 命令恢复 802.1x 在以上接口的接入控制为缺省方式	

18.2.4 （可选）配置用户认证方式

在802.1x认证中，用户通过EAP报文与设备端交互认证信息，而设备端对EAP报文采用“EAP中继”和“EAP终结”两种方式与RADIUS服务器交互认证信息，具体参见18.1.1节。

具体采用 EAP 终结还是 EAP 中继，将取决于 RADIUS 服务器的处理能力。如果RADIUS服务器的处理能力比较强，能够解析大量用户的EAP报文后再进行认证，可以采用EAP中继方式；如果RADIUS服务器处理能力不是很好，同时又需要解析大量EAP报文并完成认证，建议采用 EAP 终结方式，由设备端帮助 RADIUS 服务器完成前期的EAP解析工作。但只有采用**RADIUS**协议作为认证服务器时，**802.1x**用户的认证方式才可以配置为**EAP**中继方式 。

用户认证方式只能在系统视图下全局配置，只需在系统视图下使用 dot1x authentication-method { chap | eap |pap }命令即可。命令中的选项说明如下。

（1）chap：多选一选项，指定采用CHAP的EAP终结认证方式。

（2）eap：多选一选项，指定采用EAP中继认证方式。

（3）pap：多选一选项，指定采用PAP的EAP终结认证方式。

缺省情况下，802.1x用户认证方式为CHAP认证，可用undo dot1x authentication-method命令恢复802.1x用户的认证方式为缺省的CHAP认证。

18.2.5 （可选）使能MAC旁路认证功能

对于无法安装和无法使用 802.1x 客户端软件的终端（如打印机），可使能 MAC 旁路认证功能，使这些用户在802.1x认证过程中使用用户的MAC地址作为用户名和密码上送至认证服务器进行MAC认证（具体参见18.1.3节）。

在未使能**802.1x**功能的接口上配置了**MAC**旁路认证功能后，将会同时使能接口的**802.1x**认证功能，且首先尝试的还是**802.1x**认证，因为此时的MAC认证仅是802.1x认证失败后的一种额外认证方式选择（这也就是“旁路”和含义）。但在使能了 MAC旁路认证功能的接口下，如果想要使那些无法安装和使用802.1x客户端的终端快速地通过认证，则还可开启“旁路认证过程中优先进行MAC认证”功能，使接口优先对这些用户进行MAC认证，仅在MAC认证失败后才触发802.1x认证，以提高用户认证通过效率。

MAC 旁路认证功能可在系统视图下为多个接口进行批量配置，或在接口视图下为单个接口配置，具体步骤如表18-8所示。

表18-8 MAC旁路认证的配置步骤

步骤	命令	说明	
1	**system-view** 例如：<HUAWEI> **system-view**	进入系统视图	
2	**dot1x mac-bypass interface** { *interface-type interface-number1* [**to** *interface-number2*] } &<1-10> 例如：[HUAWEI] **dot1x mac-bypass interface** gigabitethernet 0/0/1 **to** 0/0/4	在系统视图下批量为多个接口使能 MAC 旁路认证功能。命令中的参数说明如下。 （1）{ *interface-type interface-number1* [**to** *interface-number2*] }：指定要使能 MAC 旁路认证的一个接口或者一个连续范围的多个接口 （2）&<1-10>：表示前面的{ *interface-type interface-number1* [**to** *interface-number2*] } 参数最多可以有 10 个 缺省情况下，未使能接口的 MAC 旁路认证功能，可用 **undo dot1x mac-bypass** { **interface** { *interface-type interface-number1* [**to** *interface-number2*] } &<1-10> }命令去使能指定接口的 MAC 旁路认证功能，**但此时会同时关闭指定接口的 802.1x 认证功能**	在系统视图下为一个或多个接口批量使能 MAC 旁路认证
3	**dot1x mac-bypass mac-auth-first interface** { *interface-type interface-number1* [**to** *interface-number2*] } &<1-10> 例如：[HUAWEI] **dot1x mac-bypass mac-auth-first interface** gigabitethernet 0/0/1 **to** 0/0/4	（可选）在系统视图下批量为多个接口使能旁路认证过程中优先进行 MAC 认证功能，其中的参数说明参见与第 2 步的对应参数相同 缺省情况下，未开启旁路认证过程中优先进行 MAC 认证功能，可用 **undo dot1x mac-bypass mac-auth-first interface** { *interface- type interface-number1* [**to** *interface-number2*] } &<1-10>命令在指定接口上去使能旁路认证过程中优先进行 MAC 认证功能	

（续表）

步骤	命令	说明	
4	**interface** *interface-type interface-number* 例如：[HUAWEI]**interface** gigabitethernet 0/0/1	键入要配置MAC旁路认证功能的接口，进入接口视图	在具体接口视图下为单个接口配置MAC旁路认证
5	**dot1x mac-bypass** 例如：[HUAWEI-GigabitEthernet0/0/1]**dot1x port-method port**	在以上接口上使能MAC旁路认证功能 缺省情况下，未使能接口的MAC旁路认证功能，可用 **undo dot1x mac-bypass** 命令去使能以上接口的 MAC 旁路认证功能，**但此时会同时关闭以上接口的 802.1x 认证功能**	
6	**dot1x mac-bypass mac-auth-first** 例如：[HUAWEI-GigabitEthernet0/0/1]**dot1x mac-bypass mac-auth-first**	（可选）在以上接口上使能旁路认证过程中优先进行 MAC 认证功能。缺省情况下，接口未开启旁路认证过程中优先进行 MAC 认证功能，可用 **undo dot1x mac-bypass mac-auth-first** 命令在以上接口上去使能旁路认证过程中优先进行MAC认证功能	

18.2.6 （可选）配置接口允许接入的最大802.1x认证用户数

考虑到设备性能的限制，可配置每个接口允许接入的最大802.1x认证用户数量，以限制并发访问的802.1x认证用户数。这样，当通过认证的用户数到达配置的最大数时，接口的后续用户的802.1x认证请求将直接丢弃。但在配置接口允许接入的最大用户数量之前，若该接口下的在线用户数量已经超过此最大值，此时不会影响在线用户，只会限制后续请求接入的用户。

说明

此功能只在接口的接入控制方式为基于**MAC**地址方式时有效 。当接口接入模式为基于接口方式时，接口允许接入的最大用户数量将自动设置为 1，因为此时接口只需一个用户认证成功，其他用户无需认证即可上线，具体参见本章18.2.3节。

另外，每个 S系列交换机都有一个整机并发 802.1x认证用户数的限制。S2700系列允许的最大802.1x认证并发用户数量为128，而S3700系列为512，S5700/6700系列为 256，S7700/9300/9300E/9700系列为 2048 × 接口板槽位数。在设备上各接口上配置的最大 802.1x认证用户数之和不能大于整机的最大并发用户数，否则有些接口上接入的 802.1x认证用户数虽然没有达到设置的限制，但仍不能成功通过认证。

接口允许接入的最大802.1x认证用户数可在系统视图下为多个接口进行批量配置，或在接口视图下为单个接口配置，具体步骤如表18-9所示。

表18-9 接口允许接入的最大802.1x认证用户数的配置步骤

步骤	命令	说明	
1	**system-view** 例如：<HUAWEI> **system-view**	进入系统视图	
2	**dot1x max-user** *user-number* **interface** { *interface-type interface-number1* [**to** *interface-number2*] } &<1-10> 例如：[HUAWEI] **dot1x max-user** 50 **interface** gigabitethernet 0/0/1 **to** 0/0/4	在系统视图下批量为多个接口配置允许接入的最大 802.1x 认证用户数量。命令中的参数说明如下。 （1）*user-number*：指定接口允许接入的最大 802.1x 认证用户数量。不同系列交换机的取值范围不同：除 S2700-52P-EI 、S2700-52P-PWR-EI 以外的 S2700EI 子系列为 1～8 的整数，S2700-52P-EI、S2700-52P-PWR-EI 子系列，以及 S3700/5700/6700 系列为 1～256 的整数，S7700/9300/9300E/9700 系列为 1～2 048 的整数 （2）{ *interface-type interface-number1* [**to** *interface-number2*] }：指定要配置允许接入的最大 802.1x 认证用户数量的一个接口或者一个连续范围的多个接口 （3）&<1-10>：表示前面的{ *interface-type interface-number1* [**to** *interface-number2*] } 参数最多可以有 10 个 缺省情况下，接口下允许接入的最大 802.1x 用户数量为对应系列交换机允许接入的最大用户数，可用 **undo dot1x max-user** [*user-number*] **interface** { *interface-type interface-number1* [**to** *interface-number2*] } &<1-10>命令恢复指定接口下允许接入的最大 802.1x 认证用户数量为缺省值	在系统视图下为一个或多个接口批量配置允许接入的最大 802.1x 认证用户数
3	**interface** *interface-type interface-number* 例如：[HUAWEI]**interface** gigabitethernet 0/0/1	键入要配置允许接入的最大 802.1x 认证用户数的接口，进入接口视图	在具体接口视图下为单个接口配置允许接入的最大 802.1x 认证用户数
4	**dot1x max-user** *user-number* 例如：[HUAWEI-GigabitEthernet0/0/1] **dot1x max-user** 30	配置在以上接口下允许接入的最大 802.1x 认证用户数量，参数 *user-number* 取值范围参见第 2 步该参数说明 缺省情况下，接口下允许接入的最大 802.1x 认证用户数量为对应系列交换机允许接入的最大用户数，可用 **undo dot1x max-user** [*user-number*]命令恢复以上接口下允许接入的最大 802.1x 认证用户数量为缺省值	

18.2.7 （可选）配置802.1x认证的定时器

802.1x在运行过程中会启动以下多个定时器以控制客户端、设备端和服务器端之间的报文交互（不过，一般情况下建议保持这些定时器的缺省值）。

（1）认证服务器超时定时器（server-timeout）：当设备端向认证服务器发送RADIUS Access-Request请求报文后，设备端启动此定时器。若在该定时器设置的时长内，设备端未收到认证服务器的响应，则将重发认证请求报文。

（2）客户端认证超时定时器（ client-timeout ）：当设备端向客户端发送了EAP-Request/MD5 Challenge请求报文后，设备端启动此定时器。若在该定时器设置的时长内，设备端没有收到客户端的响应，则设备端将重发该报文。

（3）用户名请求超时定时器（tx-period）：该定时器定义了两个时间间隔。一个是在客户端主动发起认证的情况下，当设备端向客户端发送EAP-Request/Identity请求报文后，设备端启动该定时器。如果在该定时器设置的时间间隔内设备端没有收到客户端的响应，则设备端将向客户端重发认证请求报文。另一个是为兼容不主动发送EAPOL-Start连接请求报文的客户端，设备端会定期以组播方式发送EAP-Request/Identity请求报文来检测客户端。

以上三种802.1x认证定时器缺省均为开启的，配置方法也很简单，只需在系统视图下使用 dot1x timer { client-timeout client-timeout-value | server-timeout server-timeout- value | tx-period tx-period-value }命令配置即可。命令中的参数说明如下。

（1）client-timeout client-timeout-value：多选一参数，配置客户端认证超时定时器，取值范围为1～120的整数秒。缺省情况下，客户端认证超时定时器为30s。

（2）server-timeout server-timeout-value：多选一参数，配置认证服务器超时定时器，取值范围为1～120的整数秒。缺省情况下，认证服务器超时定时器为30s

（3）tx-period tx-period-value：多选一参数，配置发送认证请求的超时定时器，取值范围为1～120的整数秒。缺省情况下，发送认证请求超时定时器为30s。

可分别执行以上命令为不同定时器进行设置，可用undo dot1x timer { client-timeout | server-timeout | tx-period }命令恢复 802.1x各项定时器参数为缺省值。

18.2.8 （可选）配置802.1x认证的静默功能

为了阻止那些非法用户频繁地进行认证尝试，可以在设备端使能静默功能，这样某一802.1x用户在60s内认证失败的次数超过规定的值时就会将该用户静默一段时间，阻止该用户继续尝试认证（在许多系统中都有这样的功能），在静默周期内直接丢弃该用户的802.1x认证请求。

802.1x 认证静默功能是在系统视图下全局配置的（不是针对具体接口配置的 ），具体的配置步骤如表18-10所示。

表18-10 802.1x认证静默功能的配置步骤

步骤	命令	说明
1	**system-view** 例如：<HUAWEI> **system-view**	进入系统视图
2	**dot1x quiet-period** 例如：[HUAWEI] **dot1x quiet-period**	使能静默功能。缺省情况下，未使能静默功能，可用 **undo dot1x quiet-period** 命令去使能静默定时器功能
3	**dot1x quiet-times** *fail-times* 例如：[HUAWEI] **dot1x quiet-times** 5	配置 802.1x 用户在被静默前 60s 内允许认证失败的次数，取值范围为 1～10 的整数。缺省情况下，允许认证失败的次数为 3 次，可用 **undo dot1x quiet-times** 命令恢复为缺省值
4	**dot1x timer quiet-period** *quiet-period-value* 例如：[HUAWEI] **dot1x timer quiet-period** 100	配置一旦达到第 2 步所设置的认证失败次数而对该用户静默的时长，取值范围为 10～3 600 的整数秒。缺省情况下，认证失败用户处于静默的时间为 60s，可用 **undo dot1x timer quiet-period** 命令恢复为缺省值

18.2.9 （可选）配置对802.1x认证用户进行重认证

如果管理员在认证服务器上修改了某一用户的信息，从而更改用户的访问权限、授权属性等参数。此时如果用户已经在线，则需要及时对该用户进行重认证，以确保用户的合法性。

对 802.1x 用户进行重认证有两种方式：（1）对指定接口下所有在线 802.1x 用户进行周期重认证；（2）对指定MAC地址的在线802.1x用户进行重认证，且仅进行一次重认证。

配置对802.1x用户进行重认证功能后，设备将会把保存的在线用户的认证参数发送到认证服务器进行重认证。如果认证服务器上用户的认证信息没有变化，则用户正常在线；如果用户的认证信息已更改，则该用户会被下线，然后需要用户根据更改后的认证参数重新进行接入认证。

对802.1x认证用户进行重认证可在系统视图下为多个接口进行批量配置，或在接口视图下为单个接口具体配置，具体步骤如表18-11所示。

表18-11 对802.1x用户进行重认证的配置步骤

配置方式	步骤	命令	说明	
对指定接口下所有在线802.1x用户进行周期重认证	1	**system-view** 例如：<HUAWEI> **system-view**	进入系统视图	
	2	**dot1x reauthenticate interface** { *interface-type interface-number1* [**to** *interface-number2*] } &<1-10> 例如：[HUAWEI] **dot1x reauthenticate interface** gigabitethernet 0/0/1 **to** 0/0/4	在系统视图下批量为多个接口使能对在线 802.1x 认证用户进行周期重认证功能。命令中的参数说明如下。 （1）{ *interface-type interface-number1* [**to** *interface-number2*] }：指定要使能 802.1x 周期重认证功能的一个接口或者一个连续范围的多个接口 （2）&<1-10>：表示前面的{ *interface-type interface-number1* [**to** *interface-number2*] }参数最多可以有 10 个 缺省情况下，未使能接口对在线 802.1x 认证用户进行周期重认证功能，可用 **undo dot1x reauthenticate interface** { *interface-type interface-number1* [**to** *interface-number2*] } &<1-10>命令去使能指定接口对在线 802.1x 认证用户进行周期重认证功能	在系统视图下为一个或多个接口批量配置对802.1x认证用户进行重认证
	3	**interface** *interface-type interface-number* 例如：[HUAWEI] **interface** gigabitethernet 0/0/1	键入要配置对 802.1x 用户进行重认证的接口，进入接口视图	在具体接口视图下为单个接口配置对在线802.1x认证用户进行周期重认证功能

（续表）

配置方式	步骤	命令	说明	
对指定接口下所有在线802.1x用户进行周期重认证	4	**dot1x reauthenticate** 例如：[HUAWEI-GigabitEthernet0/0/1] **dot1x reauthenticate**	使能以上接口对在线 802.1x 认证用户进行周期重认证功能。缺省情况下，未使能接口对在线 802.1x 认证用户进行周期重认证功能，可用 **undo dot1x reauthenticate** 命令去使能以上接口对在线 802.1x 认证用户进行周期重认证功能	在具体接口视图下为单个接口配置对在线802.1x认证用户进行周期重认证功能
	5	**quit** 例如：[HUAWEI-GigabitEthernet0/0/1] **quit**	退出接口视图，返回系统视图	
	6	**dot1x timer reauthenticate-period** *reauthenticate-period* 例如：[HUAWEI] **dot1x timer reauthenticate-period** 360	（可选）配置 802.1x 认证重认证周期，取值范围为 60～7 200 的整数秒。缺省情况下，802.1x 的重认证周期时间为 3600s，可用 **undo dot1x timer reauthenticate-period** 命令恢复对 802.1x 认证用户进行重认证的周期为缺省值	
对指定MAC地址的在线802.1x用户进行重认证	7	**dot1x reauthenticate mac-address** *mac-address* 例如：[HUAWEI] **dot1x reauthenticate mac-address** 00e0-fc01-0005	使能对指定 MAC 地址的在线 802.1x 认证用户进行重认证功能，**仅进行一次认证**。参数 *mac-address* 用来指定进行重认证的 802.1x 认证用户的 MAC 地址，格式为 H-H-H，其中 H 为 1 至 4 位的十六进制数 缺省情况下，未使能对指定 MAC 地址的在线 802.1x 认证用户进行重认证功能	

18.2.10 （可选）配置802.1x在线用户握手功能

为了确保设备端及时了解在线用户情况，可配置在线用户握手功能。这样设备将定期向通过802.1x认证的用户发送握手请求报文，如果用户在最大重传次数内没有应答，设备端就会主动将用户置为下线状态，以减少被这些用户保持的在线资源消耗。

注意

如果802.1x客户端不支持与设备进行握手报文的交互，则握手周期内设备将不会收到握手回应报文。因此，为了防止设备错误地认为用户下线，需要将在线用户握手功能关闭。

802.1x在线用户握手功能也是需要在系统视图下全局配置的（不是针对具体接口配置），具体的配置步骤如表18-12所示。

表18-12 802.1x在线用户握手功能的配置步骤

步骤	命令	说明
1	**system-view** 例如：<HUAWEI> **system-view**	进入系统视图
2	**dot1x handshake** 例如：[HUAWEI] **dot1x handshake**	使能设备与 802.1x 在线用户的握手功能。缺省情况下，未使能设备与802.1x在线用户的握手功能，可用 **undo dot1x handshake** 命令去使能设备与802.1x在线用户的握手功能

（续表）

步骤	命令	说明
3	**dot1x handshake packet-type** { **request-identity** | **srp-sha1-part2** } 例如：[HUAWEI] **dot1x handshake packet-type request-identity**	（可选）配置 802.1x 认证握手报文的类型（仅对配置该命令后上线的用户生效，对已经在线的用户无效）。命令中的选项说明如下。 （1）**request-identity**：二选一选项，指定 802.1x 认证握手报文的类型为 **request-identity** （2）**srp-sha1-part2**：二选一选项，指定 802.1x 认证握手报文的类型为 **srp-sha1-part2** 不同厂商市设备支持的握手报文类型不一样，缺省情况下，华为S系列交换机的802.1x认证握手报文的类型是 request-identity，可用 **undo dot1x handshake packet-type** 命令恢复为缺省类型
4	**dot1x retry** *max-retry-value* 例如：[HUAWEI] **dot1x retry** 5	（可选）配置向同一用户发送认证请求报文的最大次数，取值范围为1～10的整数。缺省情况下，向同一用户发送认证请求报文的最大次数为两次，可用 **undo dot1x retry** 命令恢复为缺省类型 【说明】本命令配置的向 802.1x 用户发送认证请求的最大次数，包括对未上线用户的发送的认证请求报文次数以及对已上线用户发送的握手请求报文次数
5	**dot1x timer handshake-period** *handshake-period-value* 例如：[HUAWEI] **dot1x timer handshake-period** 100	（可选）配置设备与 802.1x 在线用户的握手周期，取值范围为5～1 024的整数秒。缺省情况下，握手报文的发送时间间隔为 60s，可用 **undo dot1x timer handshake-period** 命令恢复为缺省值

18.2.11 （可选）配置Guest VLAN功能

有关Guest VLAN说明参见本章 18.1.1节。当Guest VLAN功能开启之后，设备允许用户在未进行 802.1x认证的情况下即可访问Guest VLAN中的资源，比如获取客户端软件，升级客户端或执行其他一些用户升级程序。

Guest VLAN功能可在系统视图下为多个接口进行批量配置，或在接口视图下为单个接口具体配置，具体步骤如表18-13所示。

表18-13 Guest VLAN功能的配置步骤

<table>
<tr><th>步骤</th><th>命令</th><th colspan="2">说明</th></tr>
<tr><td>1</td><td>system-view
例如：<HUAWEI> system-view</td><td colspan="2">进入系统视图</td></tr>
<tr><td>2</td><td>authentication guest-vlan vlan-id interface { interface-type interface-number1 [to interface-number2] } &<1-10>
例如：[HUAWEI] authentication guest-vlan 10 interface gigabitethernet 0/0/1 to 0/0/4</td><td>在系统视图下批量为多个接口配置加入的Guest VLAN。命令中的参数说明如下。
（1）vlan-id：指定以上接口要加入的Guest VLAN，取值范围为1～4 094的整数
（2）{ interface-type interface-number1 [to interface-number2] }：指定要加入指定的Guest VLAN的一个接口或者一个连续范围的多个接口
（3）&<1-10>：表示前面的{ interface-type interface-number1 [to interface-number2] }参数最多可以有10个
缺省情况下，接口下未配置要加入的Guest VLAN，可用 undo authentication guest-vlan [vlan-id] interface { interface-type interface-number1 [to interface-number2] } &<1-10>命令删除接口上配置的Guest VLAN</td><td>在系统视图下为多个接口批量配置加入的Guest VLAN</td></tr>
</table>

（续表）

<table>
<tr><th>步骤</th><th>命令</th><th colspan="2">说明</th></tr>
<tr><td>3</td><td>interface interface-type interface-number
例如：[HUAWEI]interface gigabitethernet 0/0/1</td><td>键入要配置加入Guest VLAN的接口，进入接口视图</td><td rowspan="2">在具体接口视图下为单个接口配置加入的Guest VLAN</td></tr>
<tr><td>4</td><td>authentication guest-vlanvlan-id
例如：[HUAWEI-GigabitEthernet0/0/1] authentication guest-vlan 10</td><td>配置在以上接口要加入的Guest VLAN，取值范围为1～4 094的整数。缺省情况下，接口下未配置Guest VLAN，可用 undo authentication guest-vlan [vlan-id]命令删除接口上配置的Guest VLAN</td></tr>
</table>

18.2.12 （可选）配置Restrict VLAN功能

有关Restrict VLAN说明参见本章 18.1.1节。为了满足用户在认证失败时也能够访问某些网络资源的需求，比如更新病毒库等。可在设备接口上配置Restrict VLAN，使用户在认证失败时被加入 Restrict VLAN 进而能够访问 Restrict VLAN 中的资源。**S2700/3700**系列不支持该项配置 。

注意

这里的"认证失败"是指认证服务器因某种原因而明确拒绝用户认证通过，比如 用户密码错误，而不是认证超时或网络连接断开等原因造成的认证失败。

Restrict VLAN功能可在系统视图下为多个接口进行批量配置，或在接口视图下为单个接口具体配置，具体步骤如表18-14所示。

表18-14 Restrict VLAN功能的配置步骤

步骤	命令	说明	
1	**system-view** 例如：<HUAWEI> **system-view**	进入系统视图	
2	**authentication restrict-vlan** *vlan-id* **interface** { *interface-type interface-number1* [**to** *interface-number2*] } &<1-10> 例如: [HUAWEI] **authentication restrict-vlan** 10 **interface** gigabitethernet 0/0/1 **to** 0/0/4	在系统视图下批量为多个接口配置接口加入的 Restrict VLAN。命令中的参数说明如下。 （1）*vlan-id*：指定以上接口要加入的 Restrict VLAN，取值范围为 1～4 094 的整数 （2）{ *interface-type interface-number1* [**to** *interface-number2*] }：指定要加入指定的 Restrict VLAN 的一个接口或者一个连续范围的多个接口 （3）&<1-10>：表示前面的 { *interface-type interface-number1* [**to** *interface-number2*] }参数最多可以有 10 个 缺省情况下，接口下未配置 Restrict VLAN，可用 **undo authentication restrict-vlan** [*vlan-id*] **interface** { *interface-type interface-number1* [**to** *interface-number2*] } &<1-10>命令删除接口上配置的 Restrict VLAN	在系统视图下为多个接口批量配置加入的 Restrict VLAN

（续表）

步骤	命令	说明	
3	**interface** *interface-type interface-number* 例如：[HUAWEI]**interface** gigabitethernet 0/0/1	（可选）键入要配置加入的 Restrict VLAN 的接口，进入接口视图	在具体接口视图下为单个接口配置加入的 Restrict VLAN
4	**authentication restrict-vlan** *vlan-id* 例如：[HUAWEI-GigabitEthernet0/0/1] **authentication restrict-vlan** 10	（可选）配置在以上接口要加入的 Restrict VLAN，取值范围为 1～4 094 的整数。缺省情况下，接口下未配置 Restrict VLAN，可用 **undo authentication restrict-vlan** [*vlan-id*]命令删除接口上配置的 Restrict VLAN	

18.2.13 （可选）配置Critical VLAN功能

有关Critical VLAN说明参见本章 18.1.1节。配置Critical VLAN功能之后，在接入设备与认证服务器之间的网络或者认证服务器出现故障时，802.1x认证用户将被加入Critical VLAN进而能够访问Critical VLAN中的资源。**S2700/3700**系列不支持该项配置 。

注意

只有 hybrid接口才能配置Critical VLAN，trunk接口和 access接口上配置Critical VLAN均不生效。

Critical VLAN功能可在系统视图下为多个接口进行批量配置，或在接口视图下为单个接口具体配置，具体步骤如表18-15所示。

表18-15 Critical VLAN功能的配置步骤

步骤	命令	说明	
1	**system-view** 例如：<HUAWEI> **system-view**	进入系统视图	
2	**authentication critical-vlan** *vlan-id* **interface** { *interface-type interface-number1* [**to** *interface-number2*] } &<1-10> 例如：[HUAWEI] **authentication critical-vlan** 10 **interface** gigabitethernet 0/0/1 **to** 0/0/4	在系统视图下批量为多个配置接口加入的 Critical VLAN。命令中的参数说明如下。 （1）*vlan-id*：指定以上接口要加入的Critical VLAN，取值范围为 1～4 094 的整数 （2）{ *interface-type interface-number1* [**to** *interface-number2*] }：指定要加入指定的 Critical VLAN 的一个接口或者一个连续范围的多个接口 （3）&<1-10>：表示前面的{ *interface-type interface-number1* [**to** *interface-number2*] }参数最多可以有 10 个 缺省情况下，接口下未配置 Restrict VLAN，可用 **undo authentication critical-vlan** [*vlan-id*] **interface** { *interface-type interface-number1* [**to** *interface-number2*] } &<1-10>命令删除接口上配置的 Critical VLAN	在系统视图下为多个接口批量配置加入的 Critical VLAN

（续表）

步骤	命令	说明	
3	**authentication critical eapol-success interface** { *interface-type interface-number1* [**to** *interface-number2*] } &<1-10> 例如：[HUAWEI] **authentication critical eapol-success interface** gigabitethernet 0/0/1 **to** 0/0/4	在系统视图下批量为多个接口将用户加入 Critical-VLAN 后向用户回应 Eapol-Success 报文功能。参数说明参见前面第 2 步。 缺省情况下，接口将用户加入 Critical-VLAN 后向用户回应 Eapol-Fail 报文功能，可用 **undo authentication critical eapol-success interface** { *interface-type interface-number1* [**to** *interface-number2*] } &<1-10>命令配置指定接口将用户加入 Critical-VLAN 后向用户回应 Eapol-Fail 报文功能	在系统视图下为多个接口批量配置加入的 Critical VLAN
4	**authentication max-reauth-req** *times* **interface** { *interface-type interface-number1* [**to** *interface-number2*] } &<1-10> 例如：[HUAWEI] **authentication max-reauth-req** 3 **interface** gigabitethernet 0/0/1 **to** 0/0/4	在系统视图下批量为多个接口对 Critical- Vlan 用户触发重认证的最大次数。参数 *times* 用来指定触发重认证的最大次数，取值范围为 1～20 的整数，其他参数参见前面第 2 步说明 缺省情况下，对 Critical-vlan 用户触发重认证的次数为 20，可用 **undo authentication max-reauth-req** [*times*] **interface** { *interface-type interface-number1* [**to** *interface-number2*] } &<1-10>命令恢复指定接口对 Critical-Vlan 用户触发重认证的最大次数为缺省值	
5	**interface** *interface-type interface-number* 例如：[HUAWEI]**interface** gigabitethernet 0/0/1	键入要配置加入 Critical VLAN 的接口，进入接口视图	在具体接口视图下为单个接口配置加入的 Critical VLAN
6	**authentication critical-vlan** *vlan-id* 例如：[HUAWEI-GigabitEthernet0/0/1] **authentication critical-vlan** 10	配置在以上接口要加入的 Critical VLAN，取值范围为 1～4 094 的整数。缺省情况下，接口下未配置 Restrict VLAN，可用 **undo authentication critical-vlan** [*vlan-id*]命令删除接口上配置的 Critical VLAN	
7	**authentication critical eapol-success** 例如：[HUAWEI-GigabitEthernet0/0/1]**authentication critical eapol-success**	配置以上接口将用户加入 Critical-VLAN 后向用户回应 Eapol-Success 报文功能。缺省情况下，将用户加入 Critical-VLAN 后向用户回应 Eapol-Fail 报文功能，可用 **undo authentication critical eapol-success** 命令恢复为缺省情况	
8	**authentication max-reauth-req** *times* 例如：[HUAWEI-GigabitEthernet0/0/1] **authentication max-reauth-req** 3	配置以上接口对Critical-Vlan用户触发重认证的最大次数。 缺省情况下，对 Critical-vlan 用户触发重认证的次数为 20，可用 **undo authentication max-reauth-req** 命令恢复为缺省值	

18.2.14 （可选）配置802.1x认证的接口Open功能

在 802.1x 网络部署初期（在正式使用时不要配置此项功能），管理员需要宏观地掌握网络中准备接入的用户数量、用户的认证方式、接入用户证书有效性等信息，这时可使能接口的 Open 功能。这样设备端允许该接口下的用户不需经过认证过程即可接入网络以便管理员获取其所需信息。S2700/3700系列不支持该

项配置。

说明

Open 功能仅在基于 MAC 地址接入控制方式下支持，支持正常上线用户的动态VLAN授权，不支持Guest-vlan授权。开启端口Open功能后，仅支持RADIUS远端认证，不支持本地、TACACS认证以及Portal认证。

802.1x 认证的接口 Open 功能可在系统视图下为多个接口进行批量配置，或在接口视图下为单个接口具体配置，具体步骤如表18-16所示。

表18-16 802.1x认证的接口Open功能的配置步骤

步骤	命令	说明	
1	**system-view** 例如：<HUAWEI> **system-view**	进入系统视图	
2	**authentication open interface** { *interface-type interface-number1* [**to** *interface-number2*] } &<1-10> 例如：[HUAWEI] **authentication open interface** gigabitethernet 0/0/1 **to** 0/0/4	在系统视图下批量为多个接口使能 Open 功能。 命令中的参数说明如下。 （1）{ *interface-type interface-number1* [**to** *interface-number2*] }：指定要使能 Open 功能的一个接口或者一个连续范围的多个接口 （2）&<1-10>：表示前面的{ *interface-type interface-number1* [**to** *interface-number2*] }参数最多可以有 10 个 缺省情况下，接口未使能 Open 功能，可用 **undo authentication open interface** { *interface-type interface-number1* [**to** *interface-number2*] } &<1-10>命令去使能指定接口的 Open 功能	在系统视图下为多个接口批量使能 Open 功能
3	**interface** *interface-type interface-number* 例如：[HUAWEI]**interface** gigabitethernet 0/0/1	键入要使能 Open 功能的接口，进入接口视图	在具体接口视图下为单个接口使能 Open 功能
4	**authentication open** 例如：[HUAWEI-GigabitEthernet0/0/1] **authentication open**	在以上接口下使能 Open 功能。缺省情况下，接口未使能 Open 功能，可用 **undo authentication open** 命令去使能以上接口的 Open 功能	

18.2.15 （可选）配置允许DHCP报文触发802.1x认证

在 802.1x认证网络中，如果存在用户使用PC操作系统（如微软Windows XP系统）自带的802.1x客户端，则这些用户将不能主动触发认证。这时，可为采用DHCP服务器自动IP地址的环境中配置允许DHCP报文触发802.1x认证功能。在使能允许DHCP报文触发 802.1x 认证后，设备端在收到用户的 DHCP 请求报文后即触发对用户的 802.1x认证，此时用户终端将自动弹出操作系统自带的802.1x认证界面。用户按照提示输入用户名与密码即可进行认证。

另外，通过使能 DHCP 报文触发 802.1x 认证功能，可使用户能够利用操作系统自带的802.1x客户端进行认证，认证通过后用户即可访问802.1x客户端下载界面下载并安装802.1x客户端软件，这将有助于快速部署网络。有关“快速部署”功能将在下面18.2.17节介绍。

配置允许DHCP报文触发802.1x认证的方法很简单，只需在系统视图下使用dot1x dhcp-trigger命令使能DHCP报文触发802.1x认证功能即可。缺省情况下，未使能DHCP报文触发 802.1x认证功能，可用undo dot1x dhcp-trigger命令去使能通过DHCP报文触 发对802.1x用户进行身份认证的功能。

18.2.16 （可选）配置单播报文触发802.1x认证

在 802.1x认证网络中，如果存在用户使用PC操作系统（如微软Windows XP系统）自带的802.1x客户端，则该用户将不能主动触发认证。此时可配置单播报文触发802.1x认证功能，使设备端在接收到客户端发送的 ARP 或 DHCP 请求报文时，就可以主动向该客户端发送单播认证报文，以触发认证。用户 PC 在收到

设备发来的认证报文后，会自动弹出操作系统自带的802.1x认证界面。用户按照提示输入用户名与密码即可进行认证。**S2700/3700**系列不支持该项配置 。

同样，通过使能单播报文触发802.1x认证功能，也可使这类用户能够利用操作系统自带的802.1x客户端进行认证，认证通过后用户即可访问802.1x客户端下载界面下载并安装 802.1x 客户端软件，这将有助于快速部署网络。有关“快速部署”功能将在下面18.2.17节介绍。

单播报文触发802.1x认证功能可在系统视图下为多个接口进行批量配置，或在接口视图下为单个接口具体配置，具体步骤如表18-17所示。

表18-17 单播报文触发802.1x认证功能的配置步骤

步骤	命令	说明	
1	**system-view** 例如：<HUAWEI> **system-view**	进入系统视图	
2	**dot1x unicast-trigger interface** { *interface-type interface-number1* [**to** *interface-number2*] } &<1-10> 例如：[HUAWEI] **dot1x unicast-trigger interface** gigabitethernet 0/0/1 **to** 0/0/4	在系统视图下批量为多个接口使能单播报文触发 802.1x 认证功能。命令中的参数说明如下。 （1）{ *interface-type interface-number1* [**to** *interface-number2*] }：指定要使能单播报文触发 802.1x 认证功能的一个接口或者一个连续范围的多个接口 （2）&<1-10>：表示前面的{ *interface-type interface-number1* [**to** *interface-number2*] }参数最多可以有 10 个 缺省情况下，接口未使能单播报文触发 802.1x 认证功能，可用 **undo dot1x unicast- trigger interface** { *interface-type interface-number1* [**to** *interface-number2*] } &<1-10>命令去使能指定接口的单播报文触发 802.1x 认证功能	在系统视图下为多个接口批量使能单播报文触发 802.1x 认证功能
3	**interface** *interface-type interface-number* 例如：[HUAWEI]**interface** gigabitethernet 0/0/1	键入要使能单播报文触发 802.1x 认证功能的接口，进入接口视图	在具体接口视图下为单个接口使能单播报文触发 802.1x 认证功能
4	**dot1x unicast-trigger** 例如：[HUAWEI-GigabitEthernet0/0/1]**dot1x unicast-trigger**	在以上接口下使能单播报文触发 802.1x 认证功能。缺省情况下，接口未使能单播报文触发 802.1x 认证功能，可用 **undo dot1x unicast-trigger** 命令去使能以上接口的单播报文触发 802.1x 认证功能	

18.2.17 （可选）配置802.1x快速部署功能

因为在 802.1x 认证网络部署中，每个客户端都必须安装 802.1x 客户端软件。如果要为每一个接入用户下载、升级802.1x客户端软件，则工作量可能非常大。这时，可通过为用户配置免认证 IP（free-ip）网段和用户 HTTP 访问 URL 重定向功能，实现用户802.1x客户端的快速部署。有关“快速部署功能”详情参见18.1.1节。**S2700/3700/9300E**系列不支持该项配置 。

说明

配置了免认证IP网段功能后，前面介绍的Guest VLAN、Critical VLAN以及Restrict VLAN功能将不再生效。

802.1x快速部署功能需要在系统视图下全局配置，具体的配置步骤如表18-18所示。

表18-18 802.1x快速部署功能的配置步骤

步骤	命令	说明
1	**system-view** 例如：<HUAWEI> **system-view**	进入系统视图
2	**dot1x free-ip** *ip-address* { *mask-length* | *mask-address* } 例如：[HUAWEI] **dot1x free-ip** 192.168.1.0 24	配置免认证 IP 网段，**只在接口授权状态为 auto 的情况下生效**。命令中的参数说明如下。 （1）*ip-address*：指定免认证网段的 IP 地址（是一个网络 IP 地址），点分十进制格式 （2）*mask-length*：二选一参数，指定以上免认证网段 IP 地址的子网掩码长度 （3）*mask-address*：二选一参数，指定以上免认证网段 IP 地址的子网掩码 缺省情况下，未配置免认证 IP 网段，可用 **undo dot1x free-ip** { *ip-address* { *mask-length* | *mask-address* } | **all** }命令恢复配置的免认证 IP 网段为缺省情况 **【说明】**未通过 802.1x 认证的用户在没有配置 free-ip 的情况下，不能通过 DHCP 服务器动态获得 IP 地址，但是若配置了 free-ip，用户便可以动态获得 IP 地址。当 802.1x 用户通过客户端主动下线后，为了防止恶意攻击，用户会在一段时间内受到限制而无法访问免费区网络资源 用户成为 802.1x 快速部署用户后，不能访问除 free-ip 网段以外的资源，但能够访问设备的一些资源。且免认证 IP 网段可配置多条，不同设备最多支持配置的 free-ip 网段不一样，具体参见相应产品手册
3	**dot1x url** *url-string* 例如：[HUAWEI] **dot1x url** http://www.123.com.cn	配置用户 HTTP 访问 URL 重定向功能，参数 *url-string* 用来指定重定向 URL（为 802.1x 客户端的下载页面地址，但必须 free-ip 在同一个网段内，否则无法访问指定的重定向 URL），为 1～200 个字符，区分大小写 配置好重定向 URL 后，当用户通过 IE 访问非 free-ip 网段地址时，设备会将用户访问的 URL 重定向到 802.1x 客户端下载页面，用户即从该页面下载 802.1x 客户端并进行安装 缺省情况下，未配置 802.1x 认证的重定向 URL，可用 **undo dot1x url** 命令取消配置的 802.1x 认证的重定向 URL

18.2.18 （可选）配置用户组功能

在 NAC 实际应用场景中，接入用户数量众多但用户类别却是有限的。针对这种情况，可在设备上创建用户组，并使每个用户组关联到一组ACL规则，则同一组内的用户将共用一组 ACL 规则。这时，可为用户组配置优先级以及 VLAN，不同用户组内的用户即具有了不同的优先级以及网络访问权限。这将使管理员更灵活地管理用户。

用户组功能是系统视图下全局配置的，具体的配置步骤如表18-19所示（**S2700/3700/S9300E**系列不支持）。

表18-19 用户组功能的配置步骤

步骤	命令	说明
1	**system-view** 例如：<HUAWEI> **system-view**	进入系统视图
2	**user-group** *group-name* 例如：[HUAWEI] **user-group** abc	创建用户组并进入用户组视图。参数 *group-name* 用来指定所创建的用户组名称，为 1~64 个字符，不支持空格，区分大小写 缺省情况下，未配置用户组，可用 **undo user-group** 命令删除已创建的用户组
3	**acl-id** *acl-number* 例如：[HUAWEI-user-group-abc] **acl-id** 3001	在以上用户组下绑定 ACL。参数 *acl-number* 用来指定与用户组绑定的高级 ACL 编号，取值范围为 3 000~3 999。执行该命令之前，需确保已使用命令 **acl** 与 **rule** 创建了 ACL 访问控制列表并配置了 ACL 规则 【注意】用户组下绑定的 ACL 不允许被修改，且所绑定的 ACL 总数不能超过 8 个，所有 ACL 规则总数不能超过 128 个 如果要求授权到用户组下的所有用户的网络访问权限相同，则用户组绑定的 ACL 中的规则不能配置有源 IP 地址，因为这时用户组中只有 IP 地址和该规则中的源 IP 相同的用户才能够匹配该 ACL 规则 缺省情况下，用户组下未绑定 ACL，可用 **undo acl-id** *acl-number* 命令删除用户组绑定的指定 ACL
4	**user-vlan** *vlan-id* 例如：[HUAWEI-user-group-abc] **user-vlan** 20	配置用户组 VLAN，取值范围为 1~4 094 的整数 【说明】如果需让不同用户组内的用户具有不同的网络访问权限，则可使用本命令配置用户组 VLAN。这样，当用户组内的某一用户上线后，则将被加入该用户组 VLAN 进而获取该用户组的网络访问权限 缺省情况下，未配置用户组 VLAN，可用 **undo user-vlan** 命令恢复用户组 VLAN 为缺省情况
5	**remark** { **8021p** *8021p-value* | **dscp** *dscp-value* }* 例如：[HUAWEI-user-group-abc] **remark dscp** 10	配置用户组优先级。用户组配置优先级后，用户组中的用户报文将继承该优先级，即不同的用户报文具有了不同的优先级别。这能够使管理员更加灵活地管理不同类别的用户。命令中的参数说明如下。 （1）*8021p-value*：可多选参数，指定对以太二层报文的处理优先级，取值范围是 0~7 的整数 （2）*dscp-value*：可多选参数，指定对 IP 报文的处理优先级，取值范围是 0~63 的整数 缺省情况下，未配置用户组优先级，可用 **undo remark** { **8021p** *8021p-value* | **dscp** *dscp-value* }*命令取消指定的用户组优先级

（续表）

步骤	命令	说明
6	**quit** 例如：[HUAWEI-user-group-abc] **quit**	退出用户组视图，返回系统视图
7	**user-group** *group-name* **enable** 例如：[HUAWEI] **user-group** abc **enable**	使能以上用户组的用户组功能。只有在使能用户组功能后，上面各项配置才能生效。**但使能用户组功能后，则不能修改该用户组与 ACL 的绑定关系** 缺省情况下，未使能用户组功能，可用 **undo user-group** *group-name* **enable** 命令去使能指定用户组的用户组功能

18.2.19 802.1x认证配置管理

在完成802.1x认证配置后，可执行以下display任意视图命令查看已配置的参数信息。

（1）display dot1x [statistics] [interface { interface-type interface-number1 [to interface- number2] } &<1-10>]：查看 802.1x的配置信息。

（2）display mac-address {authen |guest } [interface-type interface-number |vlanvlan-id] * [verbose]：查看系统当前存在的 authen类型或 guest类型的MAC地址表项。

（3）display user-group [group-name]：查看用户组的相关配置信息。

（4）display access-user user-groupgroup-name：查看与用户组绑定的所有用户的简要信息。

在用户视图下执行命令 reset dot1x statistics [interface { interface-type interface- number1 [to interface-

number2] }]，清除 802.1x的统计信息。

18.2.20 802.1x认证配置示例

本示例拓扑结构如图 18-9 所示，某公司内部大量用户终端通过 Switch 的 GE0/0/1接口接入网络。在该网络运行一段时间后，发现存在用户对公司内网进行攻击的现象。为确保网络的安全性，管理员对用户终端的网络访问权限采用802.1x认证方式进行控制，只有用户终端通过认证后Switch才允许其访问Internet中的资源。

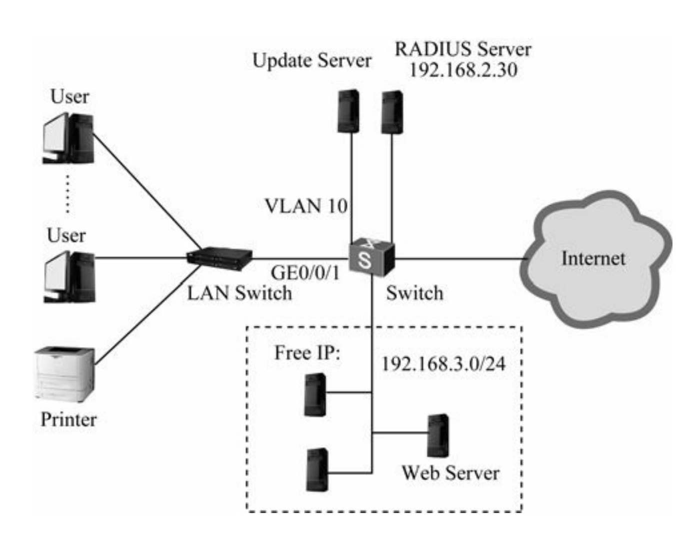


图18-9 802.1x认证配置示例拓扑结构

1．基本配置思路分析

801.1x只是一种实现网络接入控制的实现方案，具体的接入控制认证还需要通过具体的本地或者远程认证服务器来完成。本示例采用基于远程RADIUS服务器（如图中IP地址为192.168.2.30的服务器）的802.1x认证方案。因为网络中有不支持802.1x客户端的打印机设备，所以还需要在对应接口上使能MAC旁路认证。

另外，为了减轻管理员的客户端部署工作压力可选配置快速部署功能，所以在网络中专门架设了用于客户端自行下载客户端软件等资源的Web服务器（如图中的192.168.3.0/24网段中的各服务器），并设置URL重定向功能，把用户通过认证前所访问的URL重定向到以上提供免费资源的Web服务器URL上（假设为http://www.123.com.cn）。还可配置并发802.1x认证用户数限制。具体的配置思路如下（均在Switch上进行配置）。

（1）创建并配置RADIUS服务器模板、AAA方案以及ISP域，并在ISP域下绑定RADIUS服务器模板与AAA方案。保证了Switch与RADIUS服务器之间的信息交互。有关这方面的具体配置方法参见本书第17章。

（2）使能全局与接口的802.1x认证功能；使能MAC旁路认证功能，保证了无法安装和使用802.1x认证的终端（如打印机）能够通过认证。

（3）配置802.1x快速部署功能，实现用户802.1x客户端的快速部署。

（4）（可选）配置接口允许接入的最大802.1x认证用户数为200，防止过多的用户同时接入网络；配置向用户发送认证请求报文的最大次数为 3 次，防止用户不停地进行认证。

2．具体配置步骤

（1）创建并配置RADIUS服务器模板、AAA方案以及ISP域。

```
<HUAWEI>system-view
```

[HUAWEI] radius-server template rd1　#---创建RADIUS服务器模板 rd1

[HUAWEI-radius-rd1] radius-server authentication 192.168.2.30 1812　#---指定RADIUS服务器地址为192.168.2.30，服务器端口为1812

[HUAWEI-radius-rd1] radius-server shared-key cipherhello　#---指定与RADIUS服务器通信的验证密码为hello

[HUAWEI-radius-rd1] radius-server retransmit 2　#---设置RADIUS请求报文可以超时重传的次数为两次

[HUAWEI-radius-rd1] quit

[HUAWEI] aaa　#---进入AAA视图

[HUAWEI-aaa] authentication-scheme abc　#---创建AAA方案 abc

[HUAWEI-aaa-authen-abc] authentication-mode radius　#---指定采用RADIUS服务器进行认证

[HUAWEI-aaa-authen-abc] quit

[HUAWEI-aaa] domain isp1　#---创建 ISP域"isp1"

[HUAWEI-aaa-domain-isp1] authentication-scheme abc　#---指定采用名为"abc"的AAA方案

[HUAWEI-aaa-domain-isp1] radius-server rd1　#---指定采用RADIUS服务器模板 rd1

[HUAWEI-aaa-domain-isp1] quit

[HUAWEI-aaa] quit

（2）全局和接口使能802.1x认证，在接口上使能MAC旁路认证，并限制并发802.1x认证用户数为200，最多向用户发送认证请求报文3次。

[HUAWEI] dot1x enable　#---在全局下使能802.1x认证

[HUAWEI] interface gigabitethernet 0/0/1

[HUAWEI-GigabitEthernet0/0/1] dot1x enable　#---在接口下使能802.1x认证

[HUAWEI-GigabitEthernet0/0/1] dot1x mac-bypass　#---配置MAC旁路认证

[HUAWEI-GigabitEthernet0/0/1] dot1x max-user 200　#---配置接口允许接入的最大802.1x认证用户数为200

[HUAWEI-GigabitEthernet0/0/1] quit

[HUAWEI] dot1x retry 3　#---配置向用户发送认证请求报文的最大次数为3次。

（3）配置802.1x快速部署功能。

[HUAWEI] dot1x free-ip 192.168.3.0 24　!---配置免认证 IP网段，使用户在通过认证前可以访问这里面的资源

[HUAWEI] dot1x url http://www.123.com.cn　#---重定向用户访问URL

[HUAWEI] quit

配置好后，可以执行display dot1x interface任意视图命令查看 802.1x的配置信息。如下面是GE0/0/1接口上的802.1x配置信息。从中可以看到已使能802.1x的MAC旁路认证，允许接入的最大802.1x认证用户数为200等信息，符合前面的配置。

<HUAWEI>display dot1x interfacegigabitethernet0/0/1

GigabitEthernet0/0/1 status: UP　802.1x protocol is Enabled[mac-bypass]

Port control type is Auto

Authentication method is MAC-based

Reauthentication is disabled

Maximum users: 200

```
Current users: 0
There is no fast deploy user on the interface.
Guest VLAN is disabled
Critical VLAN is disabled
Restrict VLAN is disabled
Authentication Success: 0          Failure: 0
EAPOL Packets: TX    : 0           RX             : 0
Sent          EAPOL Request/Identity Packets : 0
              EAPOL Request/Challenge Packets : 0
              Multicast Trigger Packets              : 0
              EAPOL Success Packets                 : 0
              EAPOL Failure Packets                  : 0
Received      EAPOL Start Packets                    : 0
              EAPOL Logoff Packets                  : 0
              EAPOL Response/Identity Packets : 0
              EAPOL Response/Challenge Packets: 0
```

18.3 MAC认证配置与管理

MAC地址认证是一种基于端口和MAC地址对用户的网络访问权限进行控制的认证方法。它相对802.1x认证来说最大的优势是不需要用户安装任何客户端软件，用户名和密码都是用户设备的 MAC 地址。网络接入设备在首次检测到用户的 MAC 地址以后，即启动对该用户的认证。**MAC**认证主要用于对那些不支持**802.1x**客户端的设备进行接入控制，如打印机 。

与802.1x认证一样，MAC认证也只是提供了一个用户身份认证的实现方案，为了完成用户的身份认证还需要选择使用RADIUS或本地认证方法。因此也需要首先完成以下配置任务。

（1）配置用户所属的ISP认证域及其使用的AAA方案，即本地认证方案或RADIUS方案。

（2）如果需要通过RADIUS服务器进行认证，则应该在RADIUS服务器上配置相应的用户名和密码；如果需要本地认证，则应该在网络接入设备上手动添加接入用户的用户名和密码（根据**MAC**认证所采用的用户名形式进行配置 ）。

MAC认证可配置的任务如下（仅第一项为必选的，其余均为可选的 ）。

（1）使能MAC认证功能。

（2）（可选）配置用户名形式。

（3）（可选）配置用户认证域。

（4）（可选）配置接口允许接入的最大MAC认证用户数。

（5）（可选）配置MAC认证的定时器。

（6）（可选）配置对MAC认证用户进行重认证。

（7）（可选）配置Guest VLAN功能：与 18.2.11节的配置完全一样，参见即可。

（8）（可选）配置Restrict VLAN功能：与 18.2.12节的配置完全一样，参见即可。

（9）（可选）配置Critical VLAN功能：与 18.2.13节的配置完全一样，参见即可。

（10）（可选）配置用户组功能：与18.2.18节的配置完全一样，参见即可。

下面各小节分别介绍以上在本章前面未配置的各项任务的配置方法。

18.3.1 使能MAC认证功能

只有同时使能全局和接口的MAC认证功能，MAC认证的配置才能在接口上生效，具体的配置步骤如表18-20所示。但在使能**MAC**认证功能后，如果在接口下已有**MAC**认证在线用户，则不允许去使能接口的**MAC**认证功能。

表18-20 使能MAC认证功能的配置步骤

步骤	命令	说明		
1	**system-view** 例如：<HUAWEI> **system-view**	进入系统视图		
2	**mac-authen** 例如：[HUAWEI] **mac-authen**	使能全局 MAC 认证功能。缺省情况下，未使能全局 MAC 认证功能，可用 **undo mac-authen** 命令去使能设备的全局 MAC 认证功能		
3	**mac-authen interface** { *interface-type interface-number1* [**to** *interface-number2*] } &<1-10> 例如：[HUAWEI] **mac-authen interface** gigabitethernet 0/0/1 **to** 0/0/4	在系统视图下批量为多个接口使能 MAC 认证功能。命令中的参数说明如下。 （1）{ *interface-type interface-number1* [**to** *interface-number2*] }：指定要使能 MAC 认证功能的一个接口或者一个连续范围的多个接口 （2）&<1-10>：表示前面的 { *interface-type interface-number1* [**to** *interface-number2*] } 参数最多可以有 10 个 缺省情况下，所有接口都没有使能 MAC 认证功能，可用 **undo dot1x enable** [**interface** { *interface-type interface-number1* [**to** *interface- number2*] } &<1-10>]命令去使能设备指定接口上的 MAC 认证功能	在系统视图下批量使能多个接口的 MAC 认证功能	（二选一）
4	**interface** *interface-type interface-number* 例如：[HUAWEI]**interface** gigabitethernet 0/0/1	键入要使能 MAC 认证功能的接口，进入接口视图	在具体接口视图下使能 MAC 认证功能	
5	**mac-authen** 例如：[HUAWEI-GigabitEthernet0/0/1] **mac-authen**	使能以上接口的 MAC 认证功能。缺省情况下，未使能接口的 MAC 认证功能，可用 **undo mac-authen** 命令去使能以上接口的 MAC 认证功能		

注意

如果接口使能了MAC认证功能，则不能在该接口下配置如下命令，反之亦然。

（1）mac-limit：配置接口的最大MAC地址学习个数。

（2）mac-address learning disable：关闭接口的MAC地址学习功能。

（3）port link-type：配置接口的链路类型为QinQ。

（4）port vlan-mapping vlan map-vlan和port vlan-mapping vlan inner-vlan：配置接口的VLAN Mapping功能。

（5）port vlan-stacking：配置灵活QinQ功能。

（6）port-security enable：使能接口的MAC VLAN功能。

（7）mac-vlan enable：配置接口安全功能。

（8）ip-subnet-vlan enable：使能接口基于 IP子网划分VLAN的功能。

（9）port mux-vlan enable：使能接口MUX VLAN功能。

18.3.2 （可选）配置用户名形式

MAC认证用户采用的认证用户名形式有“MAC地址形式”和“固定用户名形式”两种，具体参见 18.1.3 节，可通过在系统视图下使用 mac-authen username { fixed username [password cipher password] |macaddress [format {with-hyphen |without- hyphen }] }命令进行配置。命令中的参数和选项说明如下。

● fixed username：二选一参数，指定MAC认证时使用的固定用户名，为1～64个字符串。

● cipher password：可选参数，指定以密文形式显示的MAC认证密码，可以是32位的密文密码；也可以是1～16位的明文密码。

● macaddress：二选一选项，指定以MAC地址作为MAC认证时使用的用户名。

● with-hyphen：二选一选项，指定MAC地址作为用户名输入用户名时使用带有分隔符“-”的MAC地址，例如“0005-e01c-02e3”。

● without-hyphen：二选一选项，指定MAC地址作为用户名输入用户名时使用不带有分隔符“-”的MAC地址，例如“0005e01c02e3”。

缺省情况下，MAC 认证的用户名和密码为不带分隔符“-”的 MAC 地址，可用 undo mac-authen username 命令恢复MAC地址认证时采用的用户名形式为缺省情况。

【示例 1】配置固定用户名形式认证的用户名为“vipuser”，明文格式的密码为“pass”。

<HUAWEI>system-view

[HUAWEI] mac-authen username fixed vipuserpassword cipher pass

【示例 2】配置MAC地址作为用户名进行认证，输入MAC地址时带有分隔符。

<HUAWEI>system-view

[HUAWEI] mac-authen username macaddress format with-hyphen

18.3.3 （可选）配置MAC用户认证域

当 MAC 认证用户采用的认证用户名形式为 MAC 地址形式，或者采用固定用户名形式但不带域名时，如果管理员没有配置认证域，用户将使用Default域进行认证。这会 导致众多用户在都在Default域下认证，认证方案不灵活。当MAC认证用户的用户名采用固定用户名形式，且在用户名中指定了认证域，则该用户在其自带的认证域中进行认证。

可在系统视图下为设备上所有接口全局配置或者在接口视图下为具体接口配置MAC认证用户的认证域，具体配置步骤如表18-21所示。

表18-21 MAC用户认证域的配置步骤

步骤	命令	说明
1	**system-view** 例如：<HUAWEI> **system-view**	进入系统视图
2	**mac-authen domain** *isp-name* [**mac-address** *mac-address* **mask** *mask*] 例如：[HUAWEI]**mac-authen domain** macdomain **mac-address** 1-1-1-1 ffff-ffff-ffff	配置 MAC 认证用户所使用的认证域。命令中的参数说明如下。 （1）*isp-name*：指定所在的 ISP 域名，为 1～64 个字符，不支持空格，不能使用星号“*”、问号“?”、引号“"”，不区分大小写 （2）**mac-address** *mac-address*：可选参数，指定用于 MAC 认证的用户 MAC 地址，格式为 H-H-H，其中 H 为 1 至 4 位的十六进制数 （3）**mask** *mask*：可选参数，指定以上 MAC 地址的掩码，格式为 H-H-H，其中 H 为 1 至 4 位的十六进制数，用于确定 MAC 地址的范围 缺省情况下，认证域使用系统缺省的“**default**”域，可用 **undo mac-authen domain** [*isp-name* [**mac-address** *mac-address*] | [**mac-address** { *mac-address* | **all** }]]命令取消指定或所有配置的认证域
3	**interface** *interface-type interface-number* 例如：[HUAWEI]**interface** gigabitethernet 0/0/1	（可选）键入要配置 MAC 用户认证域的接口，进入接口视图
4	**mac-authen domain** *isp-name* 例如：[HUAWEI-GigabitEthernet0/0/1]**mac-authen domain** macdomain	（可选）配置以上接口 MAC 认证用户所使用的认证域。参数用来指定对应的认证域名，其他参见第 2 步该参数说明 缺省情况下，认证域使用系统缺省的“**default**”域，可用 **undo mac-authen domain** 命令取消指定或所有配置的认证域

18.3.4 （可选）配置接口允许接入的最大**MAC**认证用户数

如果管理员需对某接口下通过 MAC 认证接入的用户数量进行限制的时候，可配置接口允许接入的 MAC 认证最大用户数量。这样，当接入用户到达配置的最大数时，后续的MAC认证用户将不能够通过该接口接入网络。

接口允许接入的最大 MAC 认证用户数可在系统视图下为多个接口进行批量配置，或在接口视图下为单个接口具体配置，具体步骤如表18-22所示。

表18-22 接口允许接入的最大MAC认证用户数的配置步骤

步骤	命令	说明	
1	**system-view** 例如：<HUAWEI> **system-view**	进入系统视图	
2	**mac-authen max-user** *user-number* **interface** { *interface-type interface-number1* [**to** *interface-number2*] } &<1-10> 例如：[HUAWEI] **mac-authen max-user** 50 **interface** gigabitethernet 0/0/1 **to** 0/0/4	在系统视图下批量为多个接口配置允许接入的最大 MAC 认证用户数量。命令中的参数说明如下。 （1）*user-number*：指定接口的最大 MAC 认证用户数量。不同系列交换机的取值范围不同：除 S2700-52P-EI、S2700-52P-PWR-EI 以外的 S2700EI 子系列为 1～8 的整数，S2700-52P-EI、S2700-52P-PWR-EI 子系列，以及 S3700/5700/6700 系列为 1～256 的整数，S7700/9300/9300E/9700 系列为 1～2048 的整数 （2）{ *interface-type interface-number1* [**to** *interface-number2*] }：指定要配置允许接入的最大 MAC 认证用户数量的一个接口或者一个连续范围的多个接口 （3）&<1-10>：表示前面的 { *interface-type interface-number1* [**to** *interface-number2*] } 参数最多可以有 10 个 缺省情况下，接口下允许接入的最大 MAC 认证用户数量为对应系列交换机允许接入的最大用户数，可用 **undo dot1x max-user** [*user-number*] **interface** { *interface-type interface-number1* [**to** *interface-number2*] } &<1-10>命令恢复为缺省值	在系统视图下为一个或多个接口批量配置允许接入的最大 MAC 认证用户数
3	**interface** *interface-type interface-number* 例如：[HUAWEI]**interface** gigabitethernet 0/0/1	键入要配置允许接入的最大 MAC 认证用户数的接口，进入接口视图	在具体接口视图下为单个接口配置允许接入的最大 MAC 认证用户数
4	**mac-authen max-user** *user-number* 例如：[HUAWEI-GigabitEthernet0/0/1] **mac-authen max-user** 30	配置在以上接口下允许接入的最大 MAC 认证用户数量，参数 *user-number* 取值范围参见第 2 步该参数说明 缺省情况下，接口下允许接入的最大 MAC 认证用户数量为对应系列交换机允许接入的最大用户数，可用 **undo dot1x max-user** [*user-number*]命令恢复为缺省值	

18.3.5 （可选）配置MAC认证定时器

MAC 认证过程中可启动多个定时器以控制接入用户、设备以及认证服务器之间进行合理、有序的交互。可配置的MAC认证定时器包括以下几种。

（1）Guest-Vlan用户重认证定时器（guest-vlan reauthenticate-period）：在用户被加入Guest Vlan之后，设备将以此定时器设置的时间间隔为周期向Guest Vlan中的用户发起重认证。若重认证成功，则用户退出Guest Vlan。

（2）用户下线探测定时器（offline-detect）：为确保用户的正常在线，设备会向在线用户发送探测报文，如果用户在探测周期内没有回应，则设备认为该用户已下线。

（3）静默定时器（quiet-period）：在用户认证失败后，设备需要静默一段时间。在静默期间，设备不处理该用户的认证请求。

（4）认证服务器超时定时器（server-timeout）：当设备向认证服务器发送 RADIUS Access-Request 请求报文后，设备启动此定时器。若在该定时器设置的时长内，设备未收到认证服务器的响应，则将重发认证请求报文。

以上这些认证定时器的配置方法是在系统视图下使用 mac-authen timer { guest- vlan reauthenticate-period interval | offline-detectoffline-detect-value |quiet-periodquiet- value | server-timeout server-timeout-value }命令进行的，命令中的参数说明如下。

（1）interval：多选一参数，指定Guest-Vlan用户重认证定时器值，取值范围为60～3 600的整数秒，缺省为 60s。

（2）offline-detect-value：多选一参数，指定用户下线探测定时器值，取值范围为30～7 200的整数秒，缺省为 300s。

（3）quiet-value：多选一参数，指定静默定时器的值，取值范围为 10～3 600的整数秒，缺省为60s。

（4）server-timeout-value：多选一参数，指定服务器超时定时器的值，取值范围为1～120的整数秒，缺省为30s。

需要分别单独为每个定时器配置，缺省情况下，guest-vlan reauthenticate-period、offline-detect、quiet-period、server-timeout定时器均为缺省开启，可用undo mac-authen timer { guest-vlan reauthenticate-period | offline-detect | quiet-period | reauthenticate- period | server-timeout }命令将指定的定时器恢复为缺省值。

18.3.6 （可选）配置对MAC认证用户进行重认证

与18.2.9节介绍的802.1x认证用户重认证一样，MAC认证中也可以为用户进行重认证。在配置了对MAC 认证用户进行重认证功能后，设备会把保存的在线用户的认证参数发送到认证服务器进行重认证，若认证服务器上用户的认证信息没有变化，则用户正常在线；若用户的认证信息已更改，则用户将会被下线，此后需要用户根据更改后的认证参数重新进行接入认证。

MAC认证中同样可以有两种重认证方式，（1）对指定接口下所有在线MAC认证用户进行周期重认证；（2）对指定 MAC 地址的在线 MAC 认证用户进行重认证，且仅进行一次重认证。

对 MAC 认证用户进行重认证可在系统视图下为多个接口进行批量配置，或在接口视图下为单个接口配置，具体步骤如表18-23所示。

表18-23 对MAC认证用户进行重认证的配置步骤

配置方式	步骤	命令	说明
对指定接口下所有在线 MAC 用户进行周期重认证	1	**system-view** 例如：<HUAWEI> **system-view**	进入系统视图

（续表）

<table>
<tr><th>配置方式</th><th>步骤</th><th>命令</th><th colspan="2">说明</th></tr>
<tr><td rowspan="5">对指定接口下所有在线 MAC 用户进行周期重认证</td><td>2</td><td>mac-authen reauthenticate interface { interface-type interface-number1 [to interface-number2] } &<1-10>
例如：[HUAWEI] mac-authen reauthenticate interface gigabitethernet 0/0/1 to 0/0/4</td><td>在系统视图下批量为多个接口使能对在线 MAC 认证用户进行周期重认证功能。
命令中的参数说明如下。
（1）{ interface-type interface-number1 [to interface-number2] }：指定要使能 MAC 周期重认证功能的一个接口或者一个连续范围的多个接口
（2）&<1-10>：表示前面的 { interface-type interface-number1 [to interface-number2] } 参数最多可以有 10 个
缺省情况下，未使能接口对在线 MAC 认证用户进行周期重认证功能，可用 undo dot1x reauthenticate interface { interface-type interface-number1 [to interface-number2] } &<1-10>命令去使能指定接口对在线MAC认证用户进行周期重认证功能</td><td>在系统视图下为一个或多个接口批量配置对 MAC 认证用户进行重认证</td></tr>
<tr><td>3</td><td>interface interface-type interface-number
例如：[HUAWEI] interface gigabitethernet 0/0/1</td><td>（可选）键入要配置对 MAC 认证用户进行重认证的接口，进入接口视图</td><td rowspan="2">在具体接口视图下为单个接口配置对在线 MAC 认证用户进行周期重认证功能</td></tr>
<tr><td>4</td><td>mac-authen reauthenticate
例如：[HUAWEI-GigabitEthernet0/0/1] mac-authen reauthenticate</td><td>（可选）使能以上接口对在线 MAC 认证用户进行周期重认证功能。缺省情况下，未使能接口对在线 MAC 认证用户进行周期重认证功能，可用 undo dot1x reauthenticate 命令去使能接口对在线 MAC 认证用户进行周期重认证功能</td></tr>
<tr><td>5</td><td>quit
例如：[HUAWEI-GigabitEthernet0/0/1] quit</td><td colspan="2">退出接口视图，返回系统视图</td></tr>
<tr><td>6</td><td>mac-authen timer reauthenticate-period reauthenticate-period
例如：[HUAWEI] mac-authen timer reauthenticate-period 360</td><td colspan="2">（可选）配置 MAC 认证重认证周期，取值范围为 60～7 200 的整数秒。缺省情况下，MAC 认证对用户的重认证周期时间为 1 800s，可用 undo mac-authen timer reauthenticate-period 命令恢复为缺省值</td></tr>
<tr><td>对指定 MAC 地址的在线 MAC 用户进行重认证</td><td>7</td><td>mac-authen reauthenticate mac-address mac-address
例如：[HUAWEI] mac-authen reauthenticate mac-address 00e0-fc01-0005</td><td colspan="2">使能对指定 MAC 地址的在线 MAC 用户进行重认证功能，仅进行一次认证。参数 mac-address 用来指定进行重认证的 MAC 用户的 MAC 地址，格式为 H-H-H，其中 H 为 1～4 位的十六进制数
缺省情况下，未使能对指定 MAC 地址的在线 MAC 认证用户进行重认证功能，可用 mac-authen reauthenticate mac-address mac-address 命令取消对指定 MAC 地址的在线 MAC 认证用户去使能重认证功能</td></tr>
</table>

18.3.7 MAC认证配置管理

在完成MAC认证配置后，可执行以下display任意视图命令查看已配置的参数信息。

（1）display mac-authen [interface { interface-type interface-number1 [to interface- number2] } &<1-10>]：查看所有或者指定接口的MAC认证配置信息。

（2）display mac-address {authen |guest } [interface-type interface-number |vlan vlan-id] *[verbose]：查看系统当前存在的 authen类型或 guest类型的MAC地址表项。

（3）display user-group [group-name]：查看所有或者指定用户组的相关配置信息。

（4）display access-user user-groupgroup-name：查看与指定用户组绑定的所有用户摘要信息。

在用户视图下执行reset mac-authen statistics [interface { interface-type interface-number1 [to interface-number2] }] 命令清除所有或者指定接口上的MAC地址认证的统计信息。

18.3.8 MAC认证配置示例

如图18-10所示，某公司内部大量打印机通过Switch的GE0/0/1接口接入网络。在该网络运行一段时间后，为增强网络的安全性，管理员需对打印机的网络访问权限进行控制。

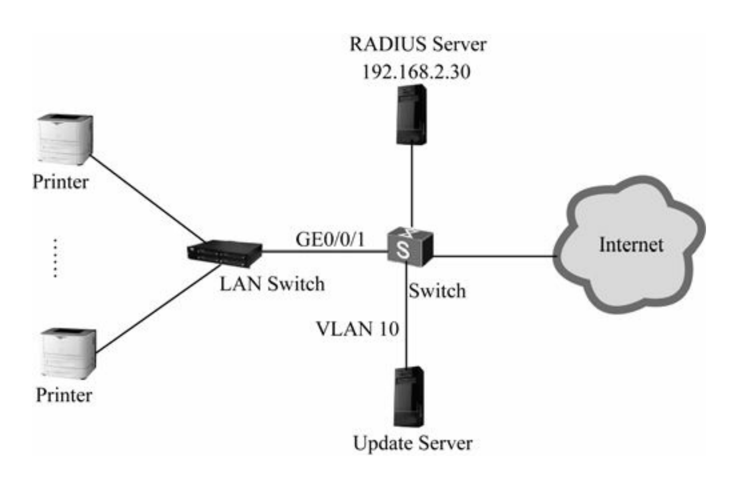


图18-10 MAC认证配置示例

1．基本配置思路分析

由于打印机无法安装和使用802.1x客户端，为实现对其网络访问权限进行限制的需求，管理员可在Switch上配置MAC认证功能。具体的配置思路如下（均在Switch上进行配置）。

（1）创建并配置RADIUS服务器模板、AAA方案以及ISP域，并在ISP域下绑定RADIUS服务器模板与AAA方案。保证了Switch与RADIUS服务器之间的信息交互。

（2）使能设备全局与接口的 MAC 认证功能。可选配置接口允许接入的最大 MAC认证用户数为 100，防止过多的用户同时接入网络。还可配置Guest VLAN为 10，满足当用户未进行认证时能够访问Guest VLAN中的资源。

2．具体的配置步骤

（1）创建并配置RADIUS服务器模板、AAA方案以及ISP域。

<HUAWEI>system-view

[HUAWEI] radius-server template rd1　#---创建并配置RADIUS服务器模板“rd1”

[HUAWEI-radius-rd1] radius-server authentication 192.168.2.30 1812　#---指定RADIUS服务器 IP地址

[HUAWEI-radius-rd1] radius-server shared-key cipherhello　#---指定与RADIUS服务器交互的共享密钥为hello

[HUAWEI-radius-rd1] radius-server retransmit 2　#---指定RADIUS服务器重传报文的次数为2

[HUAWEI-radius-rd1] quit

[HUAWEI] aaa

[HUAWEI-aaa] authentication-scheme abc　#---建AAA方案“abc”

[HUAWEI-aaa-authen-abc] authentication-mode radius　#---指定采用RADIUS服务器认证模式

[HUAWEI-aaa-authen-abc] quit

[HUAWEI-aaa] domain isp1　#---建 ISP域“isp1”

[HUAWEI-aaa-domain-isp1] authentication-scheme abc　#---绑定AAA方案“abc”

[HUAWEI-aaa-domain-isp1] radius-server rd1　#---绑定RADIUS服务器模板“rd1”

[HUAWEI-aaa-domain-isp1] quit

[HUAWEI-aaa] quit

（2）配置MAC认证。

[HUAWEI] mac-authen　#---在全局使能MAC认证

[HUAWEI] interface gigabitethernet 0/0/1

[HUAWEI-GigabitEthernet0/0/1] mac-authen　#---在接口下使能MAC认证

[HUAWEI-GigabitEthernet0/0/1] mac-authen max-user100　#---指定接口允许接入的最大MAC认证用户数为100

[HUAWEI-GigabitEthernet0/0/1] quit

[HUAWEI] vlan batch 10

[HUAWEI] authentication guest-vlan 10 interfacegigabitethernet0/0/1　#---配置MAC认证的Guest VLAN为10

配置好后，执行display mac-authen interface命令可查看接口上的MAC认证配置信息。具体如下，从中可以全面地看到前面在接口上配置的MAC认证配置信息。

[HUAWEI] display mac-authen interfacegigabitethernet0/0/1

GigabitEthernet0/0/1 state: UP.　MAC address authentication is enabled

Maximum users: 100

Current users: 0

Authentication Success: 0, Failure: 0

Guest VLAN 10 is not effective

Critical VLAN is disabled

Restrict VLAN is disabled

18.4 Portal认证配置与管理

Portal认证可使用户不经过特定客户端软件即可进行接入认证的一种基于Web页面的认证方式。Portal服务器为用户提供免费门户服务和基于Portal认证的Web页面。Portal服务器可分为外置Portal服务器与内置Portal服务器两种，具体参见18.1.4节。

与前面介绍的802.1x认证和MAC认证一样，Portal认证也只提供了一个用户身份认证的实现方案，为了完成用户的身份认证还需要选择使用RADIUS或本地认证方法。因此，也需要首先完成以下AAA配置任务（具体参见本书第17章）。

（1）配置用户所属的ISP认证域及其使用的AAA方案，即本地认证方案或RADIUS方案。

（2）如果需要通过RADIUS服务器进行认证，则应该在RADIUS服务器上配置相应的用户名和密码；如果需要本地认证，则应该在本地设备上手动添加接入用户的用户名和密码。

Portal认证的配置任务如下（仅前面两项是必选的，其他均为可选配置任务 ）。

（1）配置Portal服务器参数。

（2）使能Portal认证功能。

（3）（可选）配置与Portal服务器信息交互参数。

（4）（可选）配置Portal认证用户接入控制参数：**S2700/3700**系列不支持 。

（5）（可选）配置Portal认证用户下线探测周期：**S2700/3700**系列不支持 。

（6）（可选）配置Portal认证探测与逃生功能：**S2700/3700**系列不支持 。

（7）（可选）配置Portal认证用户信息同步功能：**S2700/3700**系列不支持 。

（8）（可选）配置Portal认证静态用户：**S2700/3700**系列不支持 。

（9）（可选）配置用户组功能：参见18.2.18节，**S2700/3700**系列不支持 。

（10）（可选）配置Portal认证静默功能：仅**S2700/3700**系列支持 。

18.4.1 配置Portal服务器参数

在Portal认证配置过程中，为保证设备与Portal服务器之间能够进行通信，必须先在设备上配置指向Portal服务器的参数，譬如指向Portal服务器的IP地址。

Portal服务器可分为外置Portal服务器与内置Portal服务器，外置Portal服务器具有独立的硬件设施，内置Portal服务器为存在于接入设备之内的内嵌实体（即由接入设备实现Portal服务器功能）。这两种Portal服务器参数的配置步骤如表18-24所示。

说明

仅S5700EI、S5700HI和S5710EI子系列支持内置Portal服务器功能。

表18-24 Portal服务器参数的配置步骤

步骤	命令	说明
1	**system-view** 例如：<HUAWEI> **system-view**	进入系统视图
配置指向外置 Portal 服务器参数		
2	**web-auth-server** *server-name* 例如：[HUAWEI] **web-auth-server** huawei	创建 Portal 服务器模板，并进入 Portal 服务器模板视图。参数 *server-name* 用来指定 Portal 服务器模板名，长度范围是 1～31 个字符，不支持空格，区分大小写 缺省情况下，未创建 Portal 服务器模板，可用 **undo web-auth-server** *server-name* 命令删除指定的 Portal 服务器模板
3	**server-ip** *server-ip-address* &<1-10> 例如：[HUAWEI-web-auth-server-huawei] **server-ip 1.1.1.1**	配置指向外置 Portal 服务器的 IP 地址，最多可配置 10 个 缺省情况下，未配置指向 Portal 服务器的 IP 地址，可用 **undo server-ip** { *server-ip-address* | **all** }命令删除指定的或者所有指向 Portal 服务器的 IP 地址
4	**url** *url-string* 例如：[HUAWEI-web-auth-server-huawei] **url** http://www.abc.com	配置指向 Portal 服务器的 URL，用于标志 Portal 认证用户可以访问的 Portal 服务器的网址，为 1～200 个字符，且必须以"http://"开头 缺省情况下，未配置指向 Portal 服务器的 URL，可用 **undo url** 命令删除指向 Portal 服务器的 URL

（续表）

步骤	命令	说明
配置指向内置 Portal 服务器参数		
2	**portal local-server ip** *ip-address* 例如：[HUAWEI] **portal local-server ip** 1.1.1.1	配置指向内置 Portal 服务器的 IP 地址 缺省情况下，未配置指向内置 Portal 服务器的 IP 地址，可用 **undo portal local-server ip** 命令删除配置指向内置 Portal 服务器的 IP 地址 【说明】配置指向内置 Portal 服务器的 IP 地址必须是设备上一个与用户之间路由可达的 IP 地址 内置 Portal 服务器的 IP 地址不能配置为设备上客户端接入网关的 IP 地址，推荐使用 LoopBack 接口地址。利用 LoopBack 接口状态稳定的优点，可以避免因为接口故障导致用户无法打开的问题。另外，由于发送到 LoopBack 接口的报文不会被转发到网络中，当请求上线的用户数目较大时，可减轻对设备性能的影响

18.4.2 使能Portal认证功能

在配置完成指向Portal服务器的参数后，设备即能够与Portal服务器进行通信。此时如果需对接入用户进行Portal认证，还必须使能设备的Portal认证功能。

针对外置Portal服务器，仅需将配置的Portal服务器模板绑定到VLANIF接口上即可对该接口下的用户进行Portal认证。而对于内置Portal服务器，则需先使能内置Portal服务器功能，然后使能设备二层接口的Portal认证功能，才能够对该接口下的用户进行Portal认证。具体的配置步骤如表18-25所示。

表18-25 使能Portal认证功能的配置步骤

步骤	命令	说明
1	**system-view** 例如：<HUAWEI> **system-view**	进入系统视图
使能外置 Portal 服务器的 Portal 认证功能		
2	**interface vlanif** *vlan-id* 例如：[HUAWEI] **interface vlanif** 10	键入要使能 Portal 认证的 VLAN 接口(与用户网络连接的 VLAN 接口)，进入 VLANIF 接口视图
3	**web-auth-server** *server-name* { **direct** | **layer3** } 例如：[HUAWEI-Vlanif10] **web-auth-server** huawei **direct**	在 VLANIF 接口下绑定 Portal 服务器模板。命令中的参数和选项说明如下。 （1）*server-name*：指定要绑定的 Portal 服务器模板，为 1～31 个字符，不支持空格，区分大小写 （2）**direct**：二选一选项，指定采用二层认证方式。当用户与设备之间没有三层转发设备时，设备能够学习到用户的 MAC 地址。此时可利用 IP 和 MAC 地址来识别用户，配置二层认证方式即可 （3）**layer3**：二选一选项，指定采用三层认证方式。当用户与设备之间存在三层转发设备时，设备不能够获取到用户的 MAC 地址，所以 IP 地址将唯一标识用户，此时需要配置为三层认证方式。

（续表）

步骤	命令	说明
3	**web-auth-server** *server-name* { **direct** | **layer3** } 例如：[HUAWEI-Vlanif10] **web-auth-server** huawei **direct**	缺省情况下，VLANIF接口下未绑定Portal服务器模板，可用 **undo web-auth-server** [*server-name* { **direct** | **layer3** }]命令删除在VLANIF接口下绑定的指定Portal服务器模板 【说明】一个VLANIF接口只能绑定一个Portal服务器模板，但同一个Portal服务器模板可绑定到不同的VLANIF接口 设备最多支持配置8个Portal服务器模板，并且最多支持在16个VLANIF接口下绑定Portal服务器模板 若在二层接口上使能了802.1x认证、MAC认证、MAC旁路认证或内置Portal认证，则不能在该接口所加入VLAN的VLANIF接口上执行本命令
使能内置Portal服务器的Portal认证功能		
2	**portal local-server https ssl-policy** *policy-name* [**port** *port-num*] 例如: [HUAWEI]**portal local-server https ssl-policy** abc	全局使能设备的内置Portal服务器功能。命令中的参数说明如下。 （1）**ssl-policy** *policy-name*：指定内置Portal服务器使用的SSL策略（此策略必须已创建），为1～23个字符，区分大小写，不支持空格 （2）**port** *port-num*：可选参数，指定https协议使用的TCP端口号，取值范围为1～65 535的整数。如果不选择此参数，则端口号为缺省值TCP 443 缺省情况下，未使能设备的内置Portal服务器功能，可用 **undo portal local-server** 命令去使能内置Portal服务器功能
3	**portal local-server enable interface** { *interface-type interface-number1* [**to** *interface-number2*] } &<1-10> 例如：[HUAWEI]**portal local-server enable interface** gigabitethernet 0/0/1	在系统视图下批量使能多个接口的Portal认证功能，但此时的接口仅可为二层接口；如果要在VLANIF接口下使能Portal认证功能，则只能使用下面第5步进行配置
4	**interface** *interface-type interface-number* 例如：[HUAWEI]**interface** gigabitethernet 0/0/1	（可选）键入要使能Portal认证功能的接口（可以是物理接口，也可以是VLANIF接口），进入接口视图
5	**portal local-server enable** 例如：[HUAWEI-GigabitEthernet0/0/1] **portal local-server enable**	（可选）在以上接口下使能Portal认证功能 缺省情况下，未使能接口的Portal认证功能，可用 **undo portal local-server enable** 命令去使能以上接口的Portal认证功能

18.4.3 （可选）配置与Portal服务器信息交互参数

如果Portal服务器为外置Portal服务器，则可通过配置设备与Portal服务器信息交互参数，以达到设备与外置Portal服务器之间正常通信同时提高信息交互安全性的目的。具体的配置步骤如表18-26所示。

表18-26 与Portal服务器信息交互参数的配置步骤

步骤	命令	说明
1	**system-view** 例如：<HUAWEI> **system-view**	进入系统视图
2	**web-auth-server version v2** [**v1**] 例如：[HUAWEI] **web-auth-server version v2**	配置设备支持的 Portal 协议版本 缺省情况下，设备同时支持 v2 与 v1 版本，建议采用缺省配置
3	**web-auth-server listening-port** *port-number* 例如：[HUAWEI] **web-auth-server listening-port** 2020	配置设备侦听 Portal 协议报文的端口号，取值范围为 1～65 535 的整数 缺省情况下，设备侦听 Portal 协议报文的端口号为 2000，可用 **web-auth-server listening-port** 命令恢复为缺省值
4	**web-auth-server reply-message** 例如：[HUAWEI] **web-auth-server reply-message**	使能将认证服务器回应的用户认证信息透传给 Portal 服务器的功能 缺省情况下，设备已使能将认证服务器回应的用户认证信息透传给 Portal 服务器的功能，可用 **undo web-auth-server reply-message** 命令去使能将认证服务器回应的用户认证信息透传给 Portal 服务器的功能
5	**web-auth-server** *server-name* 例如：[HUAWEI] **web-auth-server** huawei	键入要配置与 Portal 服务器信息交互参数的 Portal 服务器模板，进入 Portal 服务器模板视图
6	**source-ip** *ip-address* 例如：[HUAWEI-web-auth-server-huawei] **source-ip** 10.10.10.1	配置设备与 Portal 服务器通信的源 IP 地址 缺省情况下，未配置设备与 Portal 服务器通信的源 IP 地址
7	**port** *port-number* [**all**] 例如：[HUAWEI-web-auth-server-huawei] **port** 1000	配置向 Portal 服务器发送报文时使用的目的端口号。命令中的参数和选项说明如下。 （1）*port-number*：指定设备主动向认证服务器发送 UDP 报文时封装 UDP 报文的目的端口号，取值范围为 1～65 535 的整数 （2）**all**：可选项，指定设备在封装 UDP 报文时总是使用 *port-number* 参数值作为目的端口号 缺省情况下，设备向 Portal 服务器发送报文时使用的目的端口号为 50100，可用 **undo port** [**all**]命令恢复为缺省值
8	**shared-key** { **cipher** | **simple** } *key-string* 例如：[HUAWEI-web-auth-server-huawei] **shared-key simple hello**	配置设备与 Portal 服务器信息交互的共享密钥。命令中的参数说明如下。 （1）**cipher**：二选一选项，指定以密文形式显示共享密钥 （2）**simple**：二选一选项，指定以明文形式显示共享密钥 （3）*key-string*：指定共享密钥，如果选择 **simple** 选项，则为 1～16 位的明文密码；如果选择 **cipher** 选项，则可以是 32 位的密文密码，也可以是长度范围是 1～16 位明文密码 缺省情况下，未配置设备与 Portal 服务器信息交互的共享密钥，可用 **undo shared-key** 命令删除配置的共享密钥

18.4.4 （可选）配置Portal认证用户接入控制参数

在Portal认证网络的部署中，通过配置Portal认证用户的接入控制参数可以灵活地控制接入用户。譬如通过配置Portal认证用户的免认证规则可使特定的用户不经过认证或认证失败的情况下能够访问特定的网络资源；通过配置Portal认证的源认证网段，能够控制设备仅对源认证网段内的用户进行 Portal 认证，而其他网段的用户不能够通过 Portal认证接入网络。

Portal认证用户接入控制参数的配置步骤如表18-27所示。

表18-27 Portal认证用户接入控制参数的配置步骤

<table>
<tr><th>步骤</th><th>命令</th><th>说明</th></tr>
<tr><td>1</td><td>system-view
例如：<HUAWEI> system-view</td><td>进入系统视图</td></tr>
<tr><td colspan="3">配置外置 Portal 服务器的 Portal 认证用户接入控制参数</td></tr>
<tr><td>2</td><td>portal free-rule rule-id { destination { any | ip { ip-address mask { mask-length | ip-mask } | any } } |source { any | { interface interface-type interface-number | ip { ip-address mask { mask-length | ip-mask } | any } |vlan vlan-id }* } }*
或（仅 S5700HI、S5710EI 和 S6700 支持）
portal free-rule rule-id source ip ip-address mask { mask-length | ip-mask } [mac mac-address] [interfaceinterface-type interface-number] destination user-group group-name
例如：[HUAWEI] portal free-rule 0 source ip 2.2.100.0 mask 24 destination user-group static-user</td><td>配置 Portal 认证用户的免认证规则。portal 认证用户在认证成功之前，无法访问网络。通过配置免认证规则（free-rule）则可使特定的用户无需通过 portal 认证即可访问网络中的特定资源。用户免认证规则由 IP 地址、MAC 地址、所连接设备的接口与 VLAN 以及用户组等参数确定。命令中的参数和选项说明如下。
（1）rule-id：指定 Portal 认证用户免认证规则的序号，不同系列的取值范围不一样
（2）destination：指定 Portal 认证用户免认证即可访问的目的网络资源
（3）source：指定免认证即可访问目的网络资源的 Portal 认证用户来源
（4）any：指定可以是任何条件。与不同的关键字组合，具有不同的影响范围
（5）ip-address：指定 IP 地址，与不同的关键字组合，可以是指源地址或目的地址
（6）mask-length：二选一参数，指定以上 IP 地址的子网掩码长度，与不同的关键字组合，可以是指源地址掩码或目的地址掩码长度
（7）ip-mask：二选一参数，指定以上 IP 地址的子网掩码，与不同的关键字组合，可以是指源地址掩码或目的地址掩码
（8）interfaceinterface-type interface-number：指定规则中源的接口类型和接口编号，但必须属于 vlan-id 参数指定的 VLAN 中
（9）vlan-id：指定规则中源报文的 VLAN，取值范围为 1～4 094 的整数
（10）all：指定所有规则
（11）mac mac-address：指定免认证就能够访问目的网络资源的 Portal 认证用户的 MAC 地址
（12）user-group group-name：指定 Portal 认证用户允许访问的网络资源为用户组内定义的访问权限内的网络资源
使用本命令配置多条免认证规则，可累计生效，逐条匹配。缺省情况下，未配置 Portal 认证用户的免认证规则，可用 undo portal free-rule { rule-id | all }命令删除指定或所有已配置的 Portal 认证用户的免认证规则</td></tr>
<tr><td>3</td><td>portal max-user user-number
例如：[HUAWEI] portal max-user 50</td><td>配置允许接入的最大 Portal 认证用户数，不同系列的取值范围不同
缺省情况下，在规格范围内，设备允许接入的最大 Portal 认证用户数没有限制，可用 undo portal max-user 命令恢复为缺省情况</td></tr>
</table>

（续表）

步骤	命令	说明
4	**interface vlanif** *vlan-id* 例如：[HUAWEI] **interface vlanif** 10	键入要配置 Portal 认证源认证网段和强制认证域名的 VLANIF 接口，进入接口视图
5	**portal auth-network** *network-address* { *mask-length* | *mask-address* } 例如：[HUAWEI-Vlanif10] **portal auth-network** 192.168.1.0 24	在以上 VLAN 接口下配置 Portal 认证的源认证网段，也就是要对哪个网段的用户进行认证。命令中的参数说明如下。 • *network-address*：指定要进行 Portal 认证的网段 IP 地址 • { *mask-length* | *mask-address* }：指定以上网段 IP 地址的子网掩码长度或子网掩码 缺省情况下，源认证网段为 0.0.0.0/0，表示对所有网段的用户都进行认证，可用 **undo portal auth-network** { *network-address* { *mask-length* | *mask-address* } | **all** }命令恢复为缺省情况。但该命令仅对三层 Portal 认证方式有效，二层 Portal 认证方式将对所有网段的用户都进行认证
6	**portal domain** *domain-name* 例如：[HUAWEI-Vlanif10] **portal domain** abc	在以上 VLANIF 接口上配置 Portal 强制认证域名。参数 *domain-name* 用来指定 Portal 强制认证域名，为 1～64 的字符，区分大小写，不支持空格，不能使用星号"*"、问号"?"、引号""" 等特殊字符 通过配置 Portal 强制认证域名，可使所有从该接口接入的 Portal 认证用户被强制使用指定的认证域来进行认证、授权和计费。即使 Portal 用户输入的用户名中携带的域名相同，设备管理员也可以通过该配置对不同接口接入的 Portal 认证用户指定不同的认证域，从而增加了管理员部署 Portal 接入策略的灵活性 缺省情况下，未配置 Portal 强制认证域名，可用 **undo portal domain** 命令删除 Portal 强制认证域名
配置内置 Portal 服务器的 Portal 认证用户接入控制参数		
2	**portal local-server authentication-method** { **chap** | **pap** } 例如：[HUAWEI] **portal local-server authentication-method pap**	配置内置 Portal 服务器对 Portal 认证用户的认证方式。命令中的选项说明如下。 （1）**chap**：二选一选项，指定内置 Portal 服务器通过 CHAP 方式对 portal 认证用户进行认证 （2）**pap**：二选一选项，指定内置 Portal 服务器通过 PAP 方式对 portal 认证用户进行认证 缺省情况下，内置 Portal 服务器对 Portal 认证用户采用 CHAP 方式进行认证，可用 **undo portal local-server authentication-method** 命令恢复为缺省的 CHAP 认证方式

【示例 1】配置所有portal用户可以免认证访问IP地址为10.1.1.1/24的网络。

```
<HUAWEI>system-view
[HUAWEI] portal free-rule 1 destination ip10.1.1.1 mask24 source ip any
```

【示例 2】配置网段2.2.100.0/24中的实验室设备归属到用户组static-user，无需认证即可获取访问全部网络权限。

```
<HUAWEI>system-view
[HUAWEI] acl number 3100
[HUAWEI-acl-adv-3100] rule5 permit ip
[HUAWEI-acl-adv-3100] quit
[HUAWEI] user-group static-user
[HUAWEI-user-group-static-user] acl-id 3100
[HUAWEI-user-group-static-user] quit
[HUAWEI] user-group static-user enable
[HUAWEI] portal free-rule 0 source ip2.2.100.0 mask24destination user-group static-user
```

18.4.5 （可选）配置Portal认证用户下线探测周期

对于Portal认证用户，如果由于断电、网络异常断开等缘故造成用户下线，此时设备与Portal服务器上可能仍保存该用户信息，这将会造成计费不准等问题。另一方面，由于设备允许接入的用户数是有限的，若

用户异常下线但设备上仍保存用户信息，则可能导致其他用户不能接入网络。配置Portal认证用户下线探测周期后，如果用户在探测周期内没有回应，则设备认为该用户已下线。之后设备与认证服务器会及时清除该用户信息，并释放其占用的资源。但本功能仅适用于二层**Portal**认证方式，同时仅适用于外置**Portal**服务器环境。

配置 Portal 认证用户下线探测周期的方法很简单，只需在系统视图下使用 portal timer offline-detect time-length命令配置即可。取值范围为 30～7 200的整数秒，缺省情况下，下线探测周期为300s。

18.4.6 （可选）配置Portal认证探测与逃生功能

在Portal认证中，如果设备与Portal服务器之间出现网络故障导致通信中断或者Portal 服务器本身出现故障，则会造成新的 Portal 认证用户无法上线，已经在线的Portal 用户也无法正常下线。这时可配置 Portal 探测和逃生功能，使在网络故障或Portal服务器无法正常工作的情况下让用户仍然能够正常使用网络，并具有一定的网络访问权限，同时通过日志和 Trap 的方式报告故障。但本功能仅适用于外置 **Portal**服务器环境。

Portal认证探测与逃生功能的配置步骤如表18-28所示。

表18-28 Portal认证探测与逃生功能的配置步骤

步骤	命令	说明
1	**system-view** 例如：<HUAWEI> **system-view**	进入系统视图
2	**web-auth-server** *server-name* 例如: [HUAWEI] **web-auth-server** huawei	进入指定的 Portal 服务器模板视图
3	**server-detect** { **interval** *interval-period* | **max-times** *times* | **critical-num** *critical-num* | **action** { **log** | **trap** | **permit-all**} * } * 例如：[HUAWEI] **server-detect interval** 100 **max-times** 5 **critical-num** 3 **action permit-all**	使能 Portal 服务器探测与逃生功能。命令中的参数和选项说明如下： （1）**interval** *interval-period*：可多选参数，指定 Portal 服务器的探测周期，取值范围为 30～65 535 的整数秒，缺省值为 60s （2）**max-times** *times*：可多选参数，指定允许 Portal 服务器探测失败的最大次数，取值范围为 1～255 的整数，缺省值为 3 （3）**critical-num** *critical-num*：可多选参数，指定状态为 UP 的 Portal 服务器最小数目，取值范为为 0～128 的整数，缺省值为 0

（续表）

步骤	命令	说明
3	**server-detect** { **interval** *interval-period* | **max-times** *times* | **critical-num** *critical-num* | **action** { **log** | **trap** | **permit-all**} * } * 例如：[HUAWEI] **server-detect interval** 100 **max-times** 5 **critical-num** 3 **action permit-all**	（4）**action**：可多选选项，指定 Portal 服务器探测失败次数超过最大次数后的动作 （5）**log**：可多选选项，指定 Portal 服务器探测失败次数超过最大次数后发送日志信息 （6）**trap**：可多选选项，指定 Portal 服务器探测失败次数超过最大次数后发送 trap 信息 （7）**permit-all**：可多选选项，指定 Portal 服务器探测失败次数超过最大次数后取消接口的 Portal 认证功能 缺省情况下，未使能 Portal 服务器探测与逃生功能，可用 **undo server-detect** [**action** { **log** | **trap** | **permit-all** } *] 命令去使能 Portal 服务器探测与逃生功能

18.4.7 （可选）配置Portal认证用户信息同步功能

在Portal认证中，如果设备与Portal服务器之间出现网络故障导致通信中断或Portal服务器本身出现故障，使已经在线的 Portal 用户无法正常下线。这可能会导致设备与Portal 服务器用户信息不一致以及计费不准确问题。此时，可以配置用户信息同步机制来保证Portal服务器与设备上用户信息的一致性，以避免可能

出现的计费不准确问题。本功能仅适用于外置Portal服务器场景。具体的配置步骤如表18-29所示。

表18-29 Portal认证用户信息同步功能的配置步骤

步骤	命令	说明
1	**system-view** 例如：<HUAWEI> **system-view**	进入系统视图
2	**web-auth-server** *server-name* 例如：[HUAWEI] **web-auth-server** huawei	进入指定的 Portal 服务器模板视图
3	**user-sync** [**interval** *interval-period* | **max-times** *times*] * 例如：[HUAWEI]**user-sync interval** 100 **max-times** 5	使能用户信息同步功能。命令中的参数说明如下。 （1）**interval** *interval-period*：可多选参数，指定用户信息同步周期，取值范围为 30～65 535 的整数秒，缺省值为 300s （2）**max-times** *times*：可多选参数，指定用户信息同步最大失败次数，取值范围为 2～255 的整数，缺省值为 3 缺省情况下，未使能用户信息同步功能，可用 **undo user-sync** 命令去使能用户信息同步功能

18.4.8 （可选）配置Portal认证静态用户

在Portal认证中，有些终端无HTTP访问能力以致无法主动进行Portal认证。通过配置Portal认证静态用户，可使静态用户在接收到ARP报文后自动触发，使其进行Portal认证。具体的配置步骤如表18-30所示。

表18-30 Portal认证用户信息同步功能的配置步骤

步骤	命令	说明
1	**system-view** 例如：<HUAWEI> **system-view**	进入系统视图

（续表）

步骤	命令	说明
2	**static-user** *start-ip-address* [*end-ip-address*] [**vpn-instance** *vpn-instance-name*] [**domain-name** *domain-name* | **interfacein** *terface-type interface-number* [**detect**] | **mac-address** *mac-address* | **vlan** *vlan-id*] * 例如：[HUAWEI] **static-user** 1.1.1.1 1.1.1.10 **domain-name** huawei **vlan** 10	配置允许通过 ARP 报文触发 Portal 认证的静态用户。命令中的参数说明如下。（1）*start-ip-address* [*end-ip-address*]：指定静态用户所属的 IP 地址范围。如果不选择 *end-ip-address* 参数，则静态用户仅是 IP 地址为 *start-ip-address* 的用户（2）**vpn-instance** *vpn-instance-name*：可选参数，指定静态用户接入的 VPN 实例的 VPN 实例名，为 1～31 个字符，区分大小写，不支持空格（3）**domain-name** *domain-name*：可多选参数，指定静态用户进行 Portal 认证时的认证域的域名，为 1～64 个字符，不能包括空格、"*"、"？" 和 """，不能配置为 "-" 或 "—"，区分大小写（4）**interface** *interface-type interface-number*：可多选参数，指定静态用户所属接口（5）**detect**：可选项，指定允许设备主动发送 ARP 报文触发未上线静态用户进行 Portal 认证（6）**mac-address** *mac-address*：可多选参数，指定静态用户的 MAC 地址，格式为 H-H-H，其中 H 为 1 至 4 位的十六进制数（7）**vlan** *vlan-id*：可多选参数，指定静态用户所属的 VLAN，取值范围为 1～4 094 的整数 缺省情况下，未配置静态用户，可用 **undo static-user** *start-ip-address* [*end-ip-address*] [**vpn-instance** *vpn-instance-name*]命令删除指定 IP 地址范围的静态用户
3	**static-user username format-include { ip-address | mac-address | system-name }** 例如：[HUAWEI] **static-user username format-include ip-address**	配置静态用户进行 Portal 认证时使用的用户名。命令中的选项说明如下。（1）**ip-address**：多选一选项，指定静态用户的用户名为用户 IP 地址（2）**mac-address**：多选一选项，指定静态用户的用户名为用户 MAC 地址（3）**system-name**：多选一选项，指定静态用户的用户名为接入设备的设备名称。缺省情况下，静态用户的用户名为 **system-name+ip-address**，如，接入设备名称为 huawei，用户的 IP 地址为 1.1.1.1，则静态用户的用户名为：huawei1.1.1.1，可用 **undo static-user username format-include** 命令恢复静态用户的用户名为缺省情况
4	**static-user password cipher** *password* 例如：[HUAWEI]**static-user password cipher huawei**	配置静态用户进行 Portal 认证时使用的密码。参数 *password* 用来指定静态用户的密码，且以密文形式显示。取值范围可以是 1～16 位的明文密码，也可以是 32 位的密文密码，区分大小写，字符串中不能包含 "？" 和空格 缺省情况下，静态用户的密码为 vlan，可用 **undo static-user password** 命令恢复为缺省值

18.4.9 Portal认证配置管理

在完成Portal认证配置后，采用外置Portal服务器时，可使用以下display任意视图命令查看配置信息。

（1）display portal [interfacevlanif interface-number]：查看VLANIF接口下Portal认证的配置信息。

（2）display web-auth-server configuration：查看 Portal认证服务器相关的配置信息。

（3）display user-group [group-name]：查看用户组的配置信息。

（4）display access-user user-groupgroup-name：查看用户组内上线用户的信息。

（5）display static-user [domain-namedomain-name | interface interface-type interface- number1 | ip-address start-ip-address [end-ip-address] | vpn-instance vpn-instance-name]：查看静态用户的信息。

采用内置Portal服务器时，可使用以下display任意视图命令查看配置信息。

（1）display portal local-server：查看内置Portal服务器的配置信息。

（2）display portal local-server connect [user-ip ip-address]：查看内置Portal服务器上portal认证用户的连接状态。

（3）display static-user [domain-namedomain-name | interface interface-type interface- number1 | ip-address start-ip-address [end-ip-address] | vpn-instance vpn-instance-name]：查看静态用户的信息。

18.4.10 内置Portal服务器认证配置示例

本示例拓扑结构如图18-11所示，某公司内部大量用户终端通过Switch（作为接入设备）的GE0/0/1接口接入网络。在该网络运行一段时间后，发现存在用户对网络进行攻击。为确保网络的安全性，将IP地址为192.168.2.30的服务器用作RADIUS服务器，在Switch上配置内置Portal服务器认证功能（将一LoopBack接口的IP地址“192.168.1.30”配置为内置Portal服务器的IP地址）。只有用户终端通过认证后，Switch才允许其访问Internet中的资源。

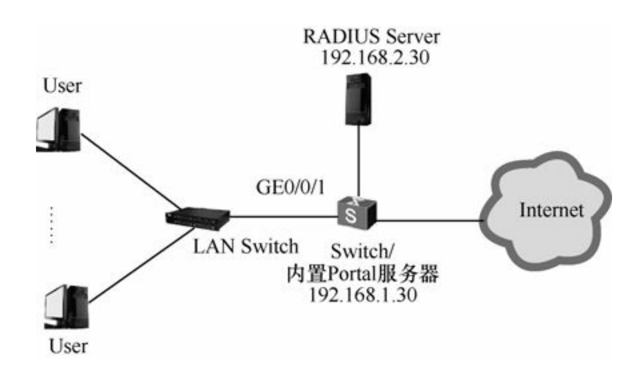


图18-11 内置Portal服务器认证配置示例拓扑结构

1．基本配置思路

根据本示例的具体要求可以得出配置思路如下（均在Switch上进行配置）。

（1）创建并配置RADIUS服务器模板、AAA方案以及ISP域，并在ISP域下绑定RADIUS服务器模板与AAA方案。保证了Switch与RADIUS服务器之间的信息交互。

（2）配置内置Portal认证，使用户终端能够通过Portal认证方式接入网络。

2．具体配置步骤

（1）创建并配置RADIUS服务器模板、AAA方案以及ISP域。这些与前面介绍的802.1x认证配置示例和MAC认证配置示例的配置方法一样，可对比18.2.20节和18.3.8节中的示例配置。

```
<HUAWEI>system-view
[HUAWEI] radius-server template rd1
[HUAWEI-radius-rd1] radius-server authentication 192.168.2.30 1812
[HUAWEI-radius-rd1] radius-server shared-key cipherhello
[HUAWEI-radius-rd1] radius-server retransmit 2
[HUAWEI-radius-rd1] quit
[HUAWEI] aaa
[HUAWEI-aaa] authentication-scheme abc
[HUAWEI-aaa-authen-abc] authentication-mode radius
[HUAWEI-aaa-authen-abc] quit
[HUAWEI-aaa] domain isp1
[HUAWEI-aaa-domain-isp1] authentication-scheme abc
[HUAWEI-aaa-domain-isp1] radius-server rd1
[HUAWEI-aaa-domain-isp1] quit
[HUAWEI-aaa] quit
```

（2）配置Portal服务器信息交互参数和HTTPS访问的SSL策略。

[HUAWEI] interface loopback 6

[HUAWEI-LoopBack6] ip address 192.168.1.30 32　#---建一个Loopback接口，并配置该Loopback接口的IP地址

[HUAWEI-LoopBack6] quit

[HUAWEI] portal local-server ip 192.168.1.30　#---置内置Portal服务器的 IP地址

[HUAWEI] ssl policy huawei

[HUAWEI-ssl-policy-huawei] certificate load pem-cert cert_rsa_cert.pem key-pair rsa key-file cert_rsa_key.pem auth-code cipher 123456@abc　#---配置打开内置Portal认证网页时所需的SSL策略

[HUAWEI-ssl-policy-huawei] quit

说明

在为SSL策略加载证书时，需确保设备上已存在所需的证书文件和密钥对文件， 否则加载不成功。有关HTTPS访问中的SSL策略配置参见本书第3章3.6.8节。

（3）在全局和接入接口上使能内置Portal认证功能。

[HUAWEI] portal local-server https ssl-policyhuawei

[HUAWEI] interface gigabitethernet 0/0/1

[HUAWEI-GigabitEthernet0/0/1] portal local-server enable

[HUAWEI-GigabitEthernet0/0/1] quit

配置好后，可通过 display portal local-server任意视图命令查看配置的内置 Portal服务器的参数信息。

<HUAWEI>display portal local-server

Portal local-server config:

server status : enable

server ip : 192.168.1.30

authentication method : chap

protocol : https

https ssl-policy : huawei

18.4.11 外置Portal服务器认证配置示例

本示例拓扑结构如图18-12所示，某公司内部大量用户终端通过Switch（作为接入 设备）的GE0/0/1接口接入网络。在该网络运行一段时间后，发现存在用户对网络进行攻击。为确保网络的安全性，将IP地址为192.168.2.30的服务器用作RADIUS服务器后，在Switch上配置Portal认证功能，并且选取的Portal服务器的IP地址为192.168.3.20，使只有用户终端通过认证后，Switch才允许其访问Internet中的资源。

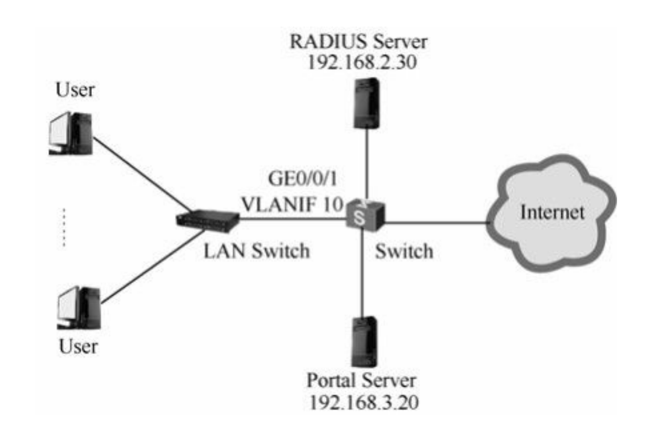


图18-12 外置Portal认证服务器认证配置示例拓扑结构

1．基本配置思路

根据本示例的具体要求可以得出配置思路如下（均在Switch上进行配置）。

（1）创建并配置RADIUS服务器模板、AAA 方案以及 ISP 域，并在 ISP 域下绑定 RADIUS 服务器模板与 AAA 方案。保证了Switch 与 RADIUS 服务器之间的信息交互。

（2）创建并配置 Portal 服务器 模板，保证设备与Portal服务器的正常信息交互。

（3）使能Portal认证功能，对接入用户进行Portal认证。

（4）（可选）配置设备与Portal服务器信息交互的共享密钥，增强设备与Portal服务器信息交互安全性。

（5）（可选）配置允许接入的最大Portal认证用户数，限制过多用户同时接入网络。

（6）（可选）配置Portal认证用户下线探测周期，保证设备能够及时删除已下线用户的信息。

（7）（可选）配置Portal认证探测与逃生功能，使用户在Portal服务器出现故障时能够正常访问网络。

2．具体配置步骤

（1）创建并配置RADIUS服务器模板、AAA方案以及ISP域。同样与前面介绍的802.1x认证配置示例、MAC认证和内置Portal服务器认证配置示例的配置方法一样。

```
<HUAWEI>system-view
[HUAWEI] radius-server template rd1
[HUAWEI-radius-rd1] radius-server authentication 192.168.2.30 1812
[HUAWEI-radius-rd1] radius-server shared-key cipherhello
[HUAWEI-radius-rd1] radius-server retransmit 2
[HUAWEI-radius-rd1] quit
[HUAWEI] aaa
[HUAWEI-aaa] authentication-scheme abc
[HUAWEI-aaa-authen-abc] authentication-mode radius
[HUAWEI-aaa-authen-abc] quit
[HUAWEI-aaa] domain isp1
[HUAWEI-aaa-domain-isp1] authentication-scheme abc
[HUAWEI-aaa-domain-isp1] radius-server rd1
[HUAWEI-aaa-domain-isp1] quit
[HUAWEI-aaa] quit
```

（2）创建并配置名称为“abc”的Portal服务器模板。

[HUAWEI] web-auth-server abc

[HUAWEI-web-auth-server-abc] server-ip 192.168.3.20

[HUAWEI-web-auth-server-abc] quit

（3）使能Portal认证功能。

[HUAWEI] interface vlanif 10

[HUAWEI-Vlanif10] web-auth-server abc direct   #---采用二层认证方式

[HUAWEI-Vlanif10] quit

（4）配置其他可选配置。

[HUAWEI] web-auth-server abc

[HUAWEI-web-auth-server-abc] shared-key cipher 12345   #---配置设备与Portal服务器信息交互的共享密钥为12345，并以密文形式显示

[HUAWEI-web-auth-server-abc] quit

[HUAWEI] portal max-user100   #---配置允许接入的最大Portal认证用户数为100

[HUAWEI] portal timer offline-detect 500   #---配置Portal认证用户下线探测周期为500秒

（5）配置Portal认证探测与逃生功能和用户信息同步功能。

[HUAWEI] web-auth-server abc

[HUAWEI-web-auth-server-abc] server-detect action log   #---使能Portal认证探测与逃生功能

[HUAWEI-web-auth-server-abc] user-sync   #---使能用户信息同步功能

[HUAWEI-web-auth-server-abc] quit

[HUAWEI] quit

配置好后，可通过 display portal任意视图命令检查在系统视图下配置的 Portal参数。

<HUAWEI>display portal

Portal timer offline-detect length:500

Portal max-user number:100

Vlanif10 protocol status: up, web-auth-server layer2(direct)

还可执行display portal interface命令查看在VLANIF接口下配置的Portal参数；执行display web-auth-server configuration命令查看Portal服务器相关的配置信息。从中可以见到以上全部的Portal认证配置信息。

<HUAWEI>display portal interface vlanif10

Vlanif10 protocol status: up, web-auth-server layer2(direct)

<HUAWEI>display web-auth-server configuration

```
  Listening port              : 2000
  Portal                      : version 1, version 2
  Include reply message       : enabled
------------------------------------------------------------------------
  Web-auth-server Name    : abc
  IP-address                  : 192.168.3.20
  Shared-key                  : %$%$qqZ$ZM:$i&] T9sF7KE～Xi%yp%$%$
  Source-IP                   : -
  Port / PortFlag             : 50100 / NO
```

```
  URL                         :
  Redirection                 : Enable
  Sync                        : Enable
  Sync Seconds                : 300
  Sync Max-times              : 3
  Detect                      : Enable
  Detect Seconds              : 60
  Detect Max-times            : 3
  Detect Critical-num         : 0
  Detect Action               : log
  Bound Vlanif                : 10
------------------------------------------------------------------------
  1 Web authentication server(s) in total
```

图书在版编目（**CIP**）数据
华为交换机学习指南/王达主编.--北京：人民邮电出版社，2014.1
（华为ICT认证系列丛书）
ISBN　978-7-115-33358-2
Ⅰ.①华…　Ⅱ.①王…　Ⅲ.①计算机网络－信息交换机－指南　Ⅳ.①TN915.05-62
中国版本图书馆CIP数据核字（2013）第239214号

内容提要

本书是国内图书市场第一本，也是目前为止唯一一本专门介绍华为交换机配置与管理的权威工具图书，同时也是华为公司指定的 ICT 认证系列培训教材。全书共 18 章，从最基础的华为 S系列园区交换机设备选型、最新5.x版本VRP系统的使用与管理，到交换机接口、以太网链接、iStack堆叠、CCS集群、各种VLAN、STP和ACL的配置与管理，再到高端应用如QoS、IP组播功能、镜像功能，以及基于MAC地址的安全管理、ARP安全管理、AAA访问控制策略、802.1x认证、MAC认证和Portal认证等各种交换机安全功能的配置与管理。

本书内容非常全面、系统，并附有大量的配置示例，同时结合了笔者20多年的工作经验，因此本书无论在专业性方面，还是在经验性和实用性方面均有更好的保障，是相关人员自学或者教学华为交换机配置与管理内容的必选教材。

♦主编　王达
责任编辑　李静
责任印制　焦志炜
♦人民邮电出版社出版发行　　北京市丰台区成寿寺路11号
邮编　100164　　电子邮件　315@ptpress.com.cn
网址　http://www.ptpress.com.cn
三河市潮河印务有限公司印刷
♦开本：787×1092　1/16
印张：60　　2014年1月第1版
字数：1422千字　　2014年1月河北第1次印刷

定价：128.00元

读者服务热线：（**010**）**81055488**　印装质量热线：（**010**）**81055316**

反盗版热线：（**010**）**81055315**